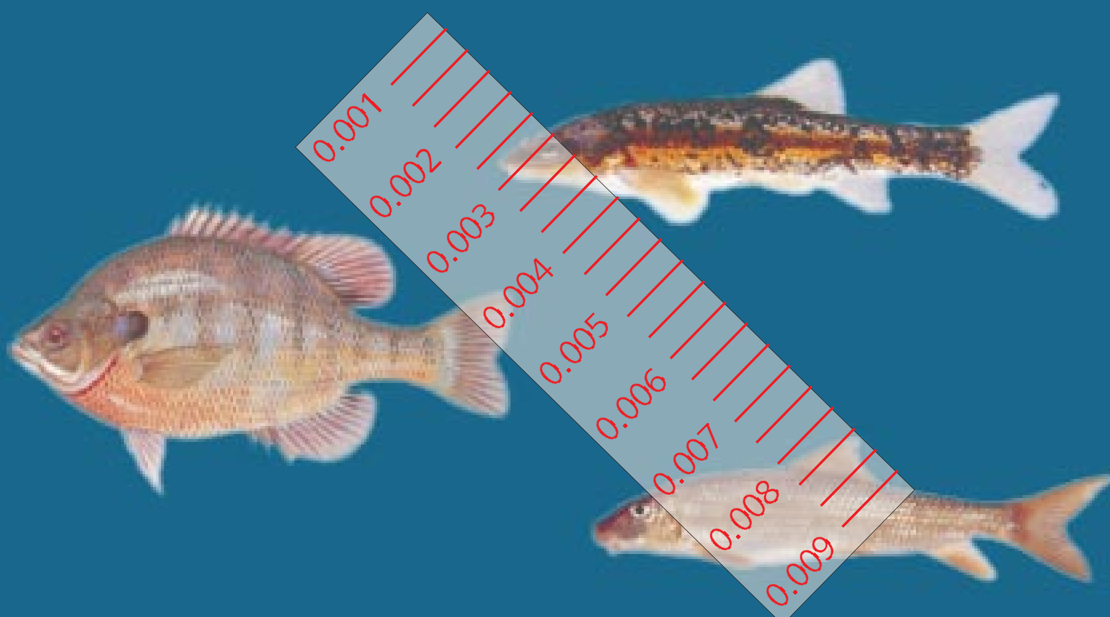
REFINEMENT AND VALIDATION OF A FISH INDE OF BIOTIC INTEGRITY FOR MARYLAND STREAMS



CHESAPEAKE BAY AND
WATERSHED PROGRAMS
MONITORING AND
NON-TIDAL ASSESSMENT

CBWP-MANTA- EA-00-

Parris N. Glendening
Governor

Kathleen K. Townsend
Lieutenant Governor

A message to Maryland's citizens

The Maryland Department of Natural Resources (DNR) seeks to preserve, protect and enhance the living resources of the state. Working in partnership with the citizens of Maryland, this worthwhile goal will become a reality. This publication provides information that will increase your understanding of how DNR strives to reach that goal through its many diverse programs.

Sarah Taylor-Rogers
Secretary

Stanley K. Arthur
Deputy Secretary

Maryland Department of Natural Resources
Tawes State Office Building
580 Taylor Avenue
Annapolis, Maryland 21401

Toll free number: 1-(877) 620 8DNR x8611
www.dnr.state.md.us

FOR FURTHER INFORMATION REGARDING THIS REPORT, PLEASE CALL 410-260-8611.
OR TOLL FREE : 1 (877) 620-8DNR x 8611

PRINTED ON RECYCLED PAPER

REFINEMENT AND VALIDATION OF A FISH INDEX OF BIOTIC INTEGRITY FOR MARYLAND STREAMS

Nancy E. Roth
Mark T. Southerland
Janis C. Chaillou
Versar, Inc.
9200 Rumsey Road
Columbia, MD 21045

and

Paul F. Kazyak
Scott A. Stranko
Maryland Department of Natural Resources
Tawes State Office Building
580 Taylor Avenue
Annapolis, MD 21401

October 2000

ABSTRACT

To aid in assessing the extent of degradation in non-tidal streams, a multi-metric Index of Biotic Integrity (IBI) based on fish assemblages was developed for the Maryland Biological Stream Survey (MBSS). The MBSS is a probability-based, statewide monitoring program designed to assess the status of biological resources and to evaluate the effects of anthropogenic activities. We used data from more than 900 MBSS sites sampled in 1994-97 to refine and validate the IBI. Three distinct geographic strata (regions), corresponding with ecoregional and physiographic boundaries, were identified via cluster analysis and multivariate analysis of variance (MANOVA) as supporting distinctly different species groups. Reference conditions were based on minimally degraded sites that were sampled in 1994-97. We quantitatively evaluated the ability of various attributes of the fish assemblage (candidate metrics) to discriminate between these reference sites and sites known to be degraded, using statistical tests and classification efficiency. Formulations of the IBI were selected for each region based on high classification efficiency and broad representation of fish assemblage attributes. Although further research is needed to develop and validate IBIs for particular habitats, including blackwater and coldwater streams, this fish IBI has proven effective in answering critical questions about the health of Maryland streams and the relative impacts of human-induced stresses on the state's aquatic systems.

TABLE OF CONTENTS

Page

F:\fibi\12337-R.wpd

LIST OF TABLES

Table No. **Page**

LIST OF FIGURES

Figure No. **Page**

1. INTRODUCTION

Biological monitoring and assessment are important components of aquatic resource management, particularly as they support the Clean Water Act goals to restore and maintain the chemical, physical, and biological integrity of surface waters. Although traditional water monitoring programs focused on chemical monitoring, water quality managers are increasingly relying on biological assessments to provide critical information on ecosystem conditions. An important advantage of using biology to characterize overall condition is that biota integrate the history of stressors at a site. One of the most widely accepted bioassessment approaches is to apply a multi-metric indicator of condition known as the Index of Biotic Integrity (IBI, Karr et al. 1986). The IBI uses the characteristics of the fish assemblage to evaluate the biological integrity at a stream site. Biological integrity is defined as the ability to support and maintain a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of the natural habitat of the region (Karr and Dudley 1981 as cited in Karr 1991). The IBI approach has been adapted for a variety of regions and taxonomic groups (e.g., Simon and Lyons 1995, Barbour et al. 1996, Weisberg et al. 1997) and is employed by a number of state water resource management programs to protect high quality resources, prevent future degradation to threatened habitats, or restore degraded sites (e.g., Yoder and Rankin 1995, USEPA 1996a).

Here, we describe the refinement and validation of an IBI for stream fishes using data from the Maryland Biological Stream Survey (MBSS). The MBSS is a multi-year, probability-based sampling program designed to assess the status of biological resources in non-tidal streams of Maryland and to determine how biological resources are affected by acidic deposition and other human activities. A provisional version of the fish IBI (Roth et al. 1998a), developed with MBSS data from the 1995 Survey and 1994 Demonstration Project, had already proven effective in evaluating selected basins in Maryland (Roth et al. 1997, 1998b). The goal of this current project was to refine this provisional IBI with the complete statewide coverage now available with the addition of 1996-1997 MBSS data. Specific objectives of this analysis included

- Refining criteria for identifying reference and degraded sites;
- Re-examining geographic stratification, to assure that regional differences in natural variability were fully addressed;
- Incorporating new information on individual species' tolerance to degradation;
- Re-testing individual metrics within each stratum to select biological measures best able to discriminate between reference and degraded conditions;
- Quantifying metric and index performance to select the most effective indicators; and
- Validating the IBI with independent data.

The refined fish IBI was a critical component of the recent comprehensive assessment of stream conditions throughout Maryland (Roth et al. 1999), based on the first complete round of MBSS sampling, 1995-1997. This statewide assessment estimated the extent of stream degradation and examined relationships between various anthropogenic stresses and the condition of biological resources. Key assessment data were provided by the fish IBI, as well as a benthic macroinvertebrate IBI (Stribling et al. 1998), a physical habitat index (Hall et al. 1999), and chemical, physical habitat, and land use parameters collected at more than 900 stream sites.

These and future applications of the fish IBI require that its development be statistically rigorous. Most recently, the Maryland Department of the Environment (MDE) is leading state efforts to employ biological criteria (biocriteria) in evaluations of surface waters, consistent with guidance issued by the U.S. Environmental Protection Agency (USEPA 1990, 1996b). Proposed protocols would incorporate MBSS fish and benthic IBI results into decisions to identify impaired waters (for listing under Section 303(d) of the Clean Water Act) and to develop Total Maximum Daily Loads (TMDLs).

To refine the IBI, we followed a series of steps previously employed in developing the provisional fish IBI for non-tidal streams in Maryland (Roth et al. 1998a). The approach involved first establishing a reference set of minimally degraded streams based on physical habitat, land use, and water quality characteristics. Multivariate analyses were used to identify regional strata supporting different species groups, thus accounting for natural variability in biological assemblages. We then compared the ability of candidate metrics (describing attributes of the fish assemblage) to discriminate between reference sites and sites known to be degraded. To aid in the selection of metrics, we tested the effectiveness of different metric combinations and analyzed the performance of the overall index. Using an independent subset of data, we validated the IBI, confirming its ability to yield effective and ecologically meaningful results. This report describes the refinement of the Maryland fish IBI, including results of metric testing and index validation, and addresses issues encountered in applying the index to ecological assessment.

2. INDICATOR REFINEMENT AND VALIDATION METHODS

Indicator development steps included developing a data base with consistent statewide coverage, identifying reference and degraded sites, determining appropriate geographic strata, testing candidate metrics, combining metrics into a multimetric index, and validating the index with independent data. A similar process was employed by members of the MBSS core team in developing two other ecological indicators for Maryland's non-tidal streams, specifically, the benthic IBI (Stribling et al. 1998) and physical habitat index (Hall et al. 1999).

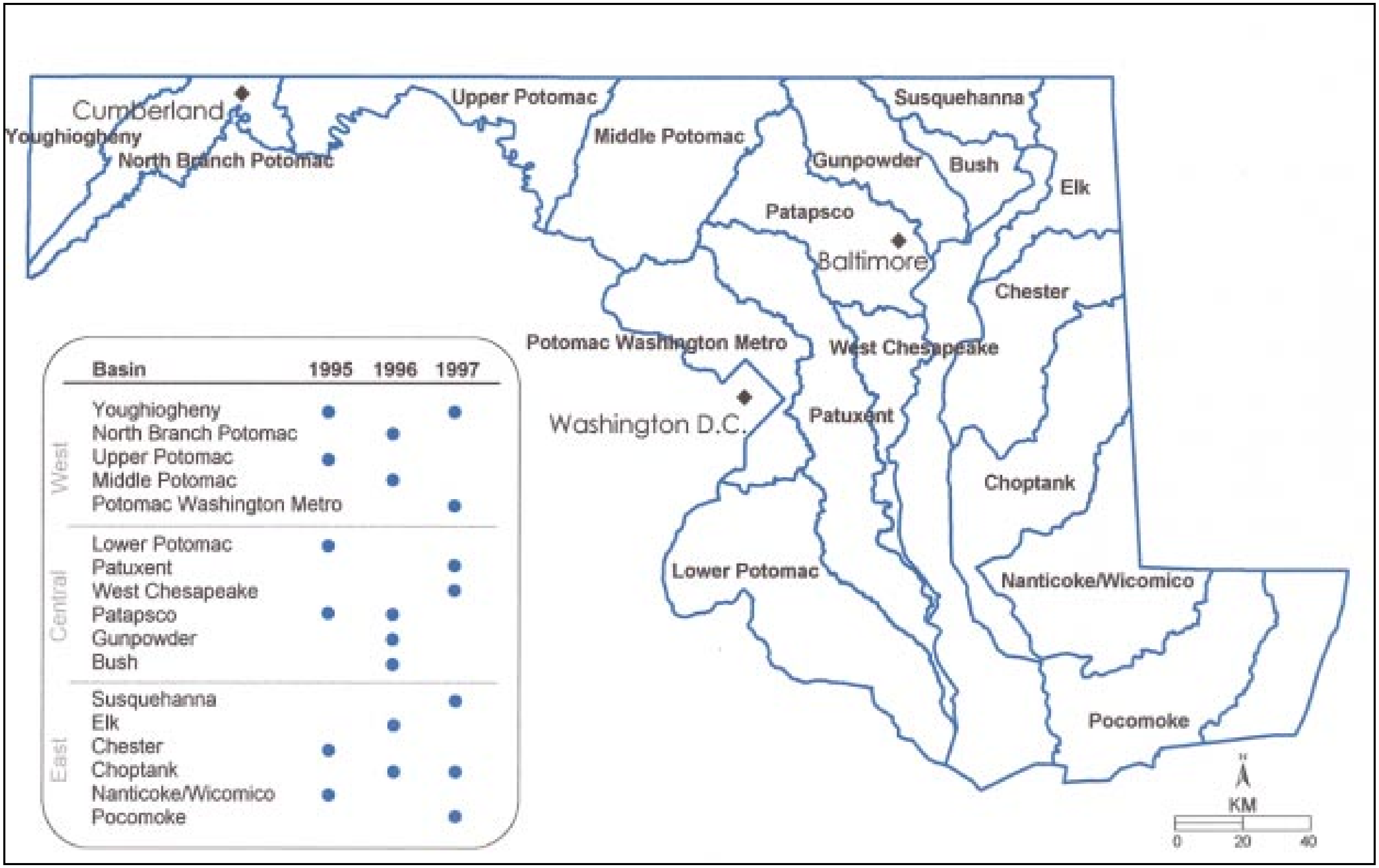
2.1 DEVELOPING THE DATA BASE

A set of 1098 stream sites sampled by the MBSS in 1994-1997 was compiled to refine and validate the fish IBI for non-tidal streams in Maryland. The statewide coverage of the data set allowed for further investigation of geographic differences than was possible for the provisional fish IBI (Roth et al. 1998a). MBSS sites, located on first- through third-order (Strahler 1957), non-tidal streams, were selected using a probability-based sampling design covering the state's major drainage basins (Figure 2-1) (Volstad et al. 1996, Roth et al. 1999). The original sampling frame for the MBSS, based on a previous statewide stream chemistry survey (the Maryland Synoptic Stream Chemistry Survey, Knapp et al. 1988), was constructed by overlaying basin boundaries on a map of all blue-line stream reaches in the state as digitized on the U.S. Geological Survey 1:250,000-scale map. Seventy-five-meter (75-m) stream segments served as the elementary sampling units for which biological, water chemistry, and physical habitat data were collected. At all sites, field data were collected using standard methods developed for the MBSS (Kazyak 1996).

Fish were sampled during the summer index period (about June 1 to September 30) using quantitative, double-pass electrofishing of the 75-m stream segments. Block nets were placed at each end of the segment and one or more direct-current, backpack electrofishing units were used to sample the entire segment. All fish captured (> 25 mm total length) were identified, counted, and weighed in aggregate; up to 100 individuals of each species were examined for external anomalies such as lesions and tumors.

During the spring index period (about March 1 to May 1), water samples were collected and pH, acid neutralizing capacity (ANC), conductivity, sulfate, nitrate, and dissolved organic carbon (DOC) were measured in the laboratory using standard methods (USEPA 1987). During summer sampling, dissolved oxygen (DO), pH, temperature, and conductivity were measured in situ.

Physical habitat assessments were conducted at all stream segments using the procedures detailed in Kazyak (1996), which were largely patterned after other widely used assessment techniques (Plafkin et al. 1989, Barbour and Stribling 1991, Ohio EPA 1987, Rankin 1989, 1995). Instream habitat structure, bank stability, degree of channel alteration, and other physical habitat



	Basin	1995	1996	1997
West	Youghiogheny	●		●
	North Branch Potomac		●	
	Upper Potomac	●		
	Middle Potomac		●	
	Potomac Washington Metro			●
Central	Lower Potomac	●		
	Patuxent			●
	West Chesapeake			●
	Patapsco	●	●	
	Gunpowder		●	
	Bush		●	
East	Susquehanna			●
	Elk		●	
	Chester	●		
	Choptank		●	●
	Nanticoke/Wicomico	●		
	Pocomoke			●

Figure 2-1. Basins in the MBSS study area and the years each were sampled in the 1995-1997 Maryland Biological Stream Survey

features were assessed qualitatively based on visual observations within each sample segment, and assigned scores (0-20 points) within four categories (optimal 16-20 points, sub-optimal 11-15, marginal 6-10, poor 0-5), following standard narrative guidelines for each category (Kazyak 1996). Observations of the surrounding area were used to assign similar ratings for aesthetic value and remoteness. Evidence of point sources, stream channelization, and other human impacts were recorded. Riparian vegetation width was estimated up to 50 m from the stream.

Land use information was extracted from Multi-Resolution Land Characteristics land cover maps in digital format (MRLC 1996). Catchments upstream from sample sites were digitized using contour lines from digital county topographic maps (1:62,500 scale) and overlaid on the land cover data. Following spatial definition of catchments, catchment areas for each site were calculated and the percentage of area as each major land use type (urban, agricultural, and forested) within each catchment was determined.

Following definition of reference and degraded sites (described below), the data sets for all four years were pooled and data were subdivided by randomly assigning sites to two groups: 2/3 of the sites were designated as "calibration" sites for refinement of the fish IBI, while the remaining 1/3 were reserved for independent validation of the IBI. Thus, both the calibration and validation sets represented conditions throughout the state and across multiple sample years.

2.2 IDENTIFYING REFERENCE AND DEGRADED SITES

Criteria for identifying reference and degraded sites for the most part followed the definitions employed by Roth et al. (1998a), with the addition of an urban land use criterion for degraded sites. Reference sites were defined as those with minimal anthropogenic disturbance, based on thresholds established for water chemistry, physical habitat, and catchment land use. Reference site criteria eliminated sites affected by extreme acidification, nutrient loading, or physical alteration. Water chemistry and riparian width thresholds were set at levels generally considered detrimental to streams (e.g., Baker et al. 1990, Osborne and Kovacic 1993). Physical habitat thresholds were set at the optimal to suboptimal level for critical parameters (Kazyak 1996). Thresholds for catchment land use, which serve as overall estimates of human influence, eliminated the most extreme cases of land use alteration. Reference sites meeting all 12 of the following criteria were identified:

- pH $\geq$ 6 or blackwater stream (pH < 6 and DOC $\geq$ 8 mg/l)
- ANC $\geq$ 50 μeq/l
- DO $\geq$ 4 ppm
- nitrate $\leq$ 300 μeq/l (4.2 mg/l)
- urban land use $\leq$ 20% of catchment area
- forest land use $\geq$ 25% of catchment area
- remoteness rating: optimal or suboptimal
- aesthetics rating: optimal or suboptimal
- instream habitat rating: optimal or suboptimal
- riparian buffer width $\geq$ 15 m

- no channelization
- no point source discharges

We next identified a set of degraded sites exhibiting any of three types of anthropogenic stress: acidification, eutrophication, or physical habitat alteration. Because new data included more sites in the Piedmont areas of central Maryland (including the Washington and Baltimore metropolitan areas), a criterion for sites in heavily urbanized catchments that lacked riparian buffer vegetation was included to more fully represent degradation from intense urban development. This criterion added 38 sites to the list of degraded sites, including 19 in the Piedmont regions. Sites meeting any of the following criteria were designated as degraded:

- pH $\leq$ 5 and ANC $\leq$ 0 μeq/l (except for blackwater streams, DOC $\geq$ 8 mg/l) (n=23 sites)
- DO $\leq$ 2 ppm (n=20)
- nitrate > 500 μeq/l (7 mg/l) and DO < 3 ppm (n=0)
- instream habitat rating poor and urban land use > 50% of catchment area (n=15)
- instream habitat rating poor and bank stability rating poor (n=34)
- instream habitat rating poor and channel alteration rating poor (n=69)
- urban land use > 50% of catchment area and riparian buffer width = 0 m (n=48)

Streams affected by physical habitat alteration were defined as those with poor instream habitat structure in combination with at least one other factor indicative of an anthropogenic source for the alteration: high degree of urban land use, poor bank stability, or indications of channel alteration. Poor instream habitat structure alone was not sufficient to designate a site as degraded, because some streams have little woody debris, boulder, or cobble substrate even under natural conditions.

In all, 152 of the original 1098 sites (14%) were designated as reference sites and 149 (14%) were designated as degraded sites, based on these physiochemical criteria (Figure 2-2).

2.3 DETERMINING APPROPRIATE STRATA

To account for the natural variation in fish assemblage composition across the large area and diverse habitats covered by the MBSS, sites were stratified into groups based on naturally occurring biological assemblages. Appropriate reference expectations could then be established separately for each group. Cluster analysis, using the Canberra metric and flexible sorting (β=-0.25) (Boesch 1977) on log-transformed percentages of species abundance, was used to identify groups of sites based on assemblage similarity. Species present at fewer than 25 sites were excluded from the analysis, so that rare species would not drive the differences seen in the clusters. Multivariate analysis of variance (MANOVA) was used to determine which groups had statistically different assemblages. Physical variables (stream order, catchment area, summer stream temperature, basin, ecoregion/subecoregion, physiographic region) associated with significantly different biological assemblages were examined to ascertain which physical variables determined assemblage composition. To ensure sufficient sample size, a total of 438 sites was used for this analysis, including all reference sites plus additional sites designated as not substantially degraded, i.e.,

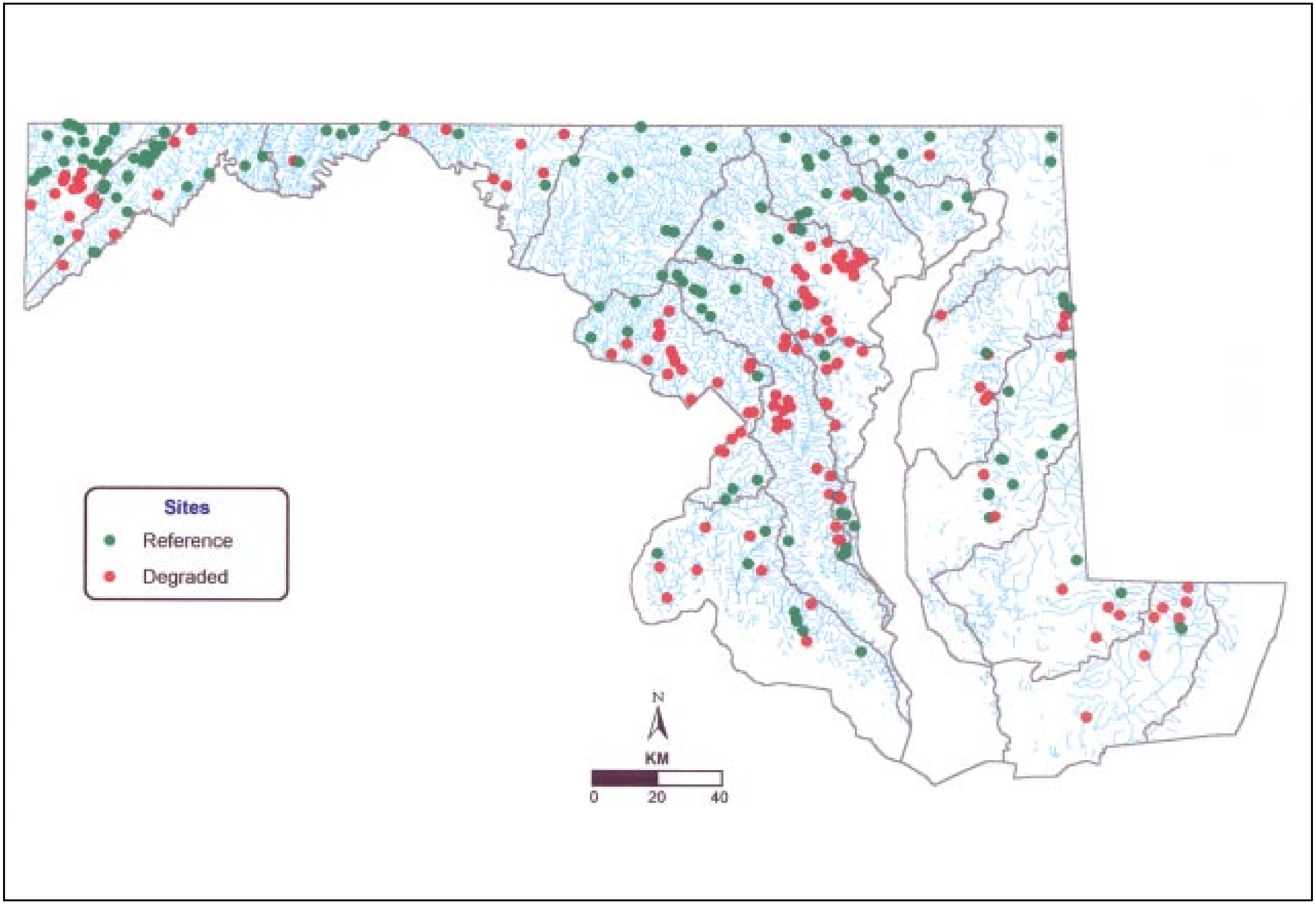


Figure 2-2. Reference and degraded sites used to refine and validate the fish Index of Biotic Integrity. Sites were selected from MBSS 1994-1997 samples.

meeting criteria slightly less stringent those used to define reference sites (e.g., pH ≥ 5.5 instead of ≥ 6, DO ≥ 3 instead of ≥ 4).

2.4 COMPILING CANDIDATE METRICS

A list of 41 candidate metrics was compiled, including multiple variations of 14 different metric types. For example, species richness was represented by two variations: total number of species and total number of native species only. The full list included each of the original 12 fish IBI metrics proposed by Karr et al. (1986) and the 9 metrics used in the provisional Maryland fish IBI (Roth et al. 1998a), along with candidate metrics from earlier IBI investigations in Maryland and other states (Hall et al. 1996a,b; Burkett and Morgan 1996; Simon and Lyons 1995). The full list of candidate metrics fell into five major groups: measures of species richness and composition, indicator species (based on tolerance), trophic function, fish abundance and condition, and reproductive function. Candidate metrics and their expected response to anthropogenic stress are listed in Table 2-1.

Fishes were classified into ecological categories, based on information in the literature, for the following characteristics: benthic fish species, trophic status, tolerance, native or introduced species, and lithophilic spawners (Table 2-2). Benthic species are those fishes that reside primarily on the stream bottom. Benthic fishes include all darters (*Etheostoma* spp., *Perca* spp.), sculpins (*Cottus* spp.), madtoms (*Noturus* spp.), and lampreys (*Petromyzon* spp., *Lampetra* spp.). Because many benthic fishes have relatively limited home ranges, they are potentially valuable indicators of local conditions. Nine trophic classifications for fishes were defined based on reported descriptions of fish diets (Jenkins and Burkhead 1993). For example, insectivores were defined as those that specialize primarily on insects, while invertivores eat insects and other invertebrates, including crustaceans, mollusks, and worms. Fishes were classified as native or introduced to the Chesapeake Bay drainage and to the Ohio River drainage (i.e., the Youghiogheny drainage in Maryland) (Lee et al. 1981, Jenkins and Burkhead 1993). Historically, streams in these two regions developed separate and distinct regional fish faunas because of their separation by the Eastern Continental Divide. Fishes reported to use rock substrates for spawning (Jenkins and Burkhead 1993) were designated as lithophilic spawners.

Tolerance of fishes to anthropogenic stress was initially determined using several literature sources (Plafkin et al. 1989, Hall et al. 1993, Hall et al. 1996a, Jacobson et al. 1992, Karr 1981, Karr et al. 1986, Lee et al. 1981, Ohio EPA 1987, EA 1993, Miller et al. 1988, Kazyak et al. 1988). Only those species that were considered clearly tolerant or intolerant by a consensus of researchers were designated as tolerant or intolerant for the "literature-based tolerance" rating using in this study; other species were designated as not rated (see Roth et al. 1997 for ratings by species).

A second rating of fish tolerance to anthropogenic stress was assigned using recent analyses of monitoring data by Maryland DNR, which ranked occurrences of individual species at degraded or minimally disturbed (reference) sites as defined for this study (S. Stranko, MDNR, unpublished data). As before, only those species that were considered clearly tolerant or intolerant were so designated (Table 2-2); species intermediate in tolerance were not rated.

Table 2-1. List of candidate metrics tested as part of refining the Maryland fish IBI and each metric's expected response to anthropogenic stress (direction of change). Metrics are categorized into five major groups. Original IBI metrics proposed for Midwestern streams (Karr et al. 1986) are indicated by an asterisk. Modifications to original metrics are grouped together.

	Metric	Expected response to stress
Species richness and composition		
1.	Total number of species *	Decrease
	Total number of native species	Decrease
2.	Number of darter species *	Decrease
	Number of darter and sculpin species	Decrease
	Number of darter, sculpin, and madtom species	Decrease
	Number of benthic fish species (darter, sculpin, madtom, and lamprey species)	Decrease
3.	Number of sunfish species *	Decrease
4.	Number of sucker species *	Decrease
	% white sucker and northern hogsucker	Increase
	% round-bodied suckers (redhorses, chubsuckers, longnose sucker, northern hogsucker)	Decrease
Indicator species		
5.	Number of intolerant species (literature-based) *	Decrease
	Number of intolerant species (based on data)	Decrease
6.	% green sunfish *	Increase
	% tolerant individuals (literature-based)	Increase
	% tolerant individuals (based on data)	Increase
	% white sucker	Increase
	% eastern mudminnow	Increase
	% creek chub	Increase
	% pioneering species	Increase
7.	% abundance of the dominant species	Increase

Table 2-1. (Continued)

Trophic composition		
8.	% omnivores *	Increase
	% generalists and omnivores	Increase
	% generalists, omnivores, and invertivores	Increase
	% omnivores and invertivores	Increase
9.	% insectivorous cyprinids *	Decrease
	% insectivores	Decrease
	% invertivores	Decrease
	% insectivores and invertivores	Decrease
10.	% top predators *	Decrease
Fish abundance and condition		
11.	Number of individuals *	Decrease
	Number of individuals per square meter	Decrease
	Number of individuals, excluding introduced species	Decrease
	Number of individuals, excluding tolerants	Decrease
12.	Biomass	Decrease
	Biomass per square meter	Decrease
13.	% of individuals with anomalies (disease, tumors, fin damage, and skeletal)* (modified as below)	Increase
	% occurrence of anomalies, including blackspot and other external parasites	Increase
	% occurrence of anomalies, excluding blackspot and other external parasites	Increase
Reproductive function		
14.	% hybrids *	Increase
	% lithophilic spawners	Decrease
	% native individuals	Decrease
	% native species	Decrease

Table 2-2. Ecological characteristics of fish species for use in IBI metrics. Tolerance: I = intolerant, T = tolerant; Native/introduced status: N = native statewide, IC = introduced to Chesapeake drainage, IY = introduced to Youghiogheny, I = introduced statewide; Trophic groups: FF = filter feeder, TP = top predator, GE = generalist, IV = invertivore, IS = insectivore, OM = omnivore, AL = algivore, HE = herbivore; NOTYPE = no category assigned.

Common Name	Tolerance (Based on Data)	Native or Introduced	Trophic Status	Lithophilic Spawner
LAMPREY (UNKNOWN)	NOTYPE	N	FF	N
LAMPREY SP.	NOTYPE	N	FF	N
AMERICAN BROOK LAMPREY	NOTYPE	N	FF	N
LEAST BROOK LAMPREY	NOTYPE	N	FF	N
SEA LAMPREY	I	N	FF	N
LONGNOSE GAR	NOTYPE	N	TP	N
AMERICAN EEL	NOTYPE	N	GE	N
BLUEBACK HERRING	NOTYPE	N	IV	N
GIZZARD SHAD	NOTYPE	N	FF	N
CHAIN PICKEREL	NOTYPE	IY	TP	N
NORTHERN PIKE	NOTYPE	IC	TP	N
REDFIN PICKEREL	T	IY	TP	N
EASTERN MUDMINNOW	T	N	IV	N
CYPRINID HYBRID	NOTYPE	NOTYPE	NOTYPE	NOTYPE
CYPRINID (UNKNOWN)	NOTYPE	NOTYPE	NOTYPE	NOTYPE
BLACKNOSE DACE	T	N	OM	N
BLUNTNOSE MINNOW	T	N	OM	N
CENTRAL STONEROLLER	I	N	AL	Y
COMELY SHINER	I	N	IV	Y
COMMON CARP	NOTYPE	I	OM	N
COMMON SHINER	I	N	OM	Y
CREEK CHUB	T	N	GE	Y
CUTLIPS MINNOW	I	N	IV	Y
CYPRINELLA SP.	NOTYPE	N	IV	N
EASTERN SILVERY MINNOW	NOTYPE	N	AL	N
FALLFISH	I	N	GE	Y
FATHEAD MINNOW	NOTYPE	I	OM	N
GOLDEN SHINER	T	N	OM	N
GOLDFISH	NOTYPE	I	OM	N
IRONCOLOR SHINER	I	N	IS	Y
LONGNOSE DACE	I	N	OM	N
LUXILUS SP.	NOTYPE	N	OM	Y
NOTROPIS SP.	NOTYPE	NOTYPE	NOTYPE	NOTYPE
PEARL DACE	NOTYPE	N	IV	Y
RIVER CHUB	I	N	OM	Y
ROSYFACE SHINER	NOTYPE	N	IV	Y
ROSYSIDE DACE	I	N	IV	Y
SATINFIN SHINER	I	N	IV	N
SILVERJAW MINNOW	NOTYPE	N	OM	Y

Table 2-2. (Continued)

Common Name	Tolerance (Based on Data)	Native or Introduced	Trophic Status	Lithophilic Spawner
SPOTFIN SHINER	I	N	IV	N
SPOTTAIL SHINER	I	N	OM	Y
STRIPED SHINER	I	N	OM	Y
SWALLOWTAIL SHINER	I	N	IV	Y
CREEK CHUBSUCKER	NOTYPE	N	IV	N
GOLDEN REDHORSE	NOTYPE	N	OM	Y
NORTHERN HOGSUCKER	I	N	IV	Y
SHORTHEAD REDHORSE	NOTYPE	N	OM	Y
WHITE SUCKER	T	N	OM	Y
BULLHEAD (UNKNOWN)	NOTYPE	N	OM	N
BROWN BULLHEAD	T	N	OM	N
CHANNEL CATFISH	NOTYPE	IC	OM	N
MARGINED MADTOM	I	IY	IV	N
TADPOLE MADTOM	NOTYPE	N	IV	N
WHITE CATFISH	NOTYPE	IY	OM	N
YELLOW BULLHEAD	NOTYPE	N	OM	N
BROOK TROUT	I	N	GE	Y
BROWN TROUT	NOTYPE	I	TP	Y
CUTTHROAT TROUT	NOTYPE	I	TP	Y
RAINBOW TROUT	NOTYPE	I	TP	Y
PIRATE PERCH	T	N	IV	N
BANDED KILLIFISH	NOTYPE	N	IV	N
MUMMICHOG	NOTYPE	N	IV	N
RAINWATER KILLIFISH	NOTYPE	N	IV	N
MOSQUITOFISH	NOTYPE	N	IV	N
SCULPIN (UNKNOWN)	NOTYPE	N	IS	Y
CHECKERED SCULPIN	NOTYPE	N	IS	Y
MOTTLED SCULPIN	I	N	IS	Y
POTOMAC SCULPIN	NOTYPE	N	IS	Y
STRIPED BASS	NOTYPE	N	TP	N
WHITE PERCH	NOTYPE	N	IV	N
SUNFISH (UNKNOWN)	NOTYPE	NOTYPE	NOTYPE	NOTYPE
BANDED SUNFISH	NOTYPE	N	IV	N
BLACK CRAPPIE	NOTYPE	IC	GE	N
BLUEGILL	T	IC	IV	N
BLUESPOTTED SUNFISH	NOTYPE	N	IV	N
FLIER	NOTYPE	N	IV	N
GREEN SUNFISH	T	IC	GE	N
LARGEMOUTH BASS	T	IC	TP	N
LONGEAR SUNFISH	NOTYPE	IC	IV	Y
MUD SUNFISH	I	N	IV	N
PUMPKINSEED	T	IY	IV	N
REDBREAST SUNFISH	I	IY	GE	N
ROCK BASS	NOTYPE	IC	GE	Y
SMALLMOUTH BASS	NOTYPE	IC	TP	N
WARMOUTH	NOTYPE	N	GE	N
LEPOMIS HYBRID	NOTYPE	NOTYPE	NOTYPE	NOTYPE
DARTER (UNKNOWN)	NOTYPE	N	NOTYPE	Y
BANDED DARTER	NOTYPE	I	IS	Y
FANTAIL DARTER	NOTYPE	N	IS	Y

Table 2-2. (Continued)

Common Name	Tolerance (Based on Data)	Native or Introduced	Trophic Status	Lithophilic Spawner
GLASSY DARTER	NOTYPE	N	IS	Y
GREENSIDE DARTER	NOTYPE	N	IS	N
JOHNNY DARTER	NOTYPE	N	IV	N
LOGPERCH	NOTYPE	N	IV	Y
RAINBOW DARTER	NOTYPE	N	IS	Y
SHIELD DARTER	I	N	IS	Y
STRIPEBACK DARTER	NOTYPE	N	IV	N
SWAMP DARTER	NOTYPE	N	IV	N
TESSELLATED DARTER	T	N	IV	N
YELLOW PERCH	NOTYPE	IY	GE	N
SPOT	NOTYPE	N	IV	N

In addition, two metrics characterizing the health or condition of individual fish were developed. The percent occurrence of all anomaly types was calculated as the number of visible anomalies per fish examined; a second variation excluded blackspot and other visible external parasites.

Because some characteristics of fish assemblages, such as abundance or species richness, tend to vary with stream size, reference expectations and IBI metric scoring should account for this natural variability (Karr et al. 1986). To account for this variation, some IBIs use separate scoring criteria for each stream order; others adjust scores by catchment area (Ohio EPA 1987). To evaluate the potential effect of stream size on each of our metrics, we examined plots of each metric against log of catchment area for each stratum, using the reference sites plus other sites not substantially degraded, from the calibration data set. We then adjusted the metrics exhibiting a strong relationship with catchment area in all three regions, as determined by significant Spearman correlations ($p < 0.05$ in at least two regions), strong linear relationship ($r^2 > 0.25$ in at least one region), and appearance of the plot. Adjusted values for metrics were calculated using the following equation, with values of m (slope) and b (intercept) derived from regression analyses:

adjusted value = observed value / expected value
where expected value = m * log (catchment area in acres) + b

The MBSS data set included a number of very small headwater streams, with correspondingly low species richness and fish abundance. If the expected total number of fish at a site (predicted for all streams of that size) fell below 100 or the expected number of fish species fell below 5, previous research (Roth et al. 1998a) had shown it to be practically impossible to characterize a reference condition accurately. For each stratum, log-linear relationships with catchment area indicated that the expected values of total abundance and number of species were below these minimum thresholds when the catchment area was less than 300 acres. Therefore, all sites on very small headwater streams (catchment area < 300 acres) were excluded from further analyses, reducing the total number of sites in the data set from 1098 to 964. Given these exclusions, a total of 105 reference sites and 81 degraded sites were available for the calibration step, with an additional 40 reference and 36 degraded sites available for indicator validation (Table 2-3).

2.5 TESTING CANDIDATE METRICS

Using the calibration data set, the 41 metrics were evaluated using statistical tests, graphical analyses, and classification efficiencies. Two statistical tests were used to compare metric values at reference sites with those at degraded sites. The Mann-Whitney U test was used to test for differences in median; the distributions of reference and degraded site values were compared using the Kolmogorov-Smirnov test. Metrics were evaluated separately for each of the three strata. Metrics showing significant differences in both of these initial statistical tests were retained for subsequent analyses.

Graphical plots comparing raw metric scores at reference and degraded sites were examined to provide further insight into metric performance. Several metrics with extremely low values (at both reference and degraded sites) were identified in this manner and eliminated from further consideration.

Table 2-3. Number of sites in fish IBI recalibration and validation data sets (watersheds ≥ 300 acres)

	Coastal Plain	Eastern Piedmont	Highlands	Total
Calibration				
Reference	33	27	45	105
Degraded	41	15	25	81
All sites	231	202	243	676
Validation				
Reference	10	11	19	40
Degraded	25	6	5	36
All sites	103	89	96	288

Metrics with low values at both reference and degraded sites generally described biological attributes that were not well represented in the samples (e.g., occurring in < 1% of the assemblage); these were not considered robust measures of condition, as values could be highly influenced by the presence of a few individuals.

The scoring of IBI metrics was based on the distribution of values observed at reference sites within each stratum. The IBI approach involves scoring each metric as 5, 3, or 1, depending on whether its value at a site approximates, deviates slightly from, or deviates greatly from conditions at the best reference sites (Karr et al. 1986). In other IBI applications (e.g., Fore et al. 1996, Lyons et al. 1996, Barbour et al. 1996), a number of different methods have been used to establish scoring thresholds, based on varying subdivisions of observed values. For our analysis, threshold values for each selected metric were established as approximately the 10th and 50th (median) percentile values for reference sites (Figure 2-3) and were established separately for each stratum. For each metric expected to decrease with degradation, values below the 10th percentile were scored as 1. Values between the 10th and 50th percentiles were scored as 3, as they fell short of median expected values for reference sites. Values above the 50th percentile were scored as 5. Scoring was reversed for metrics expected to increase with degradation (e.g., values below the 50th percentile were scored as 5, and values above the 90th percentile were scored as 1). This method differs from other scoring systems in that both the upper and lower thresholds are independently derived from the distribution of reference site values. The 10 percentile threshold for designating scores of 1 represents an intent

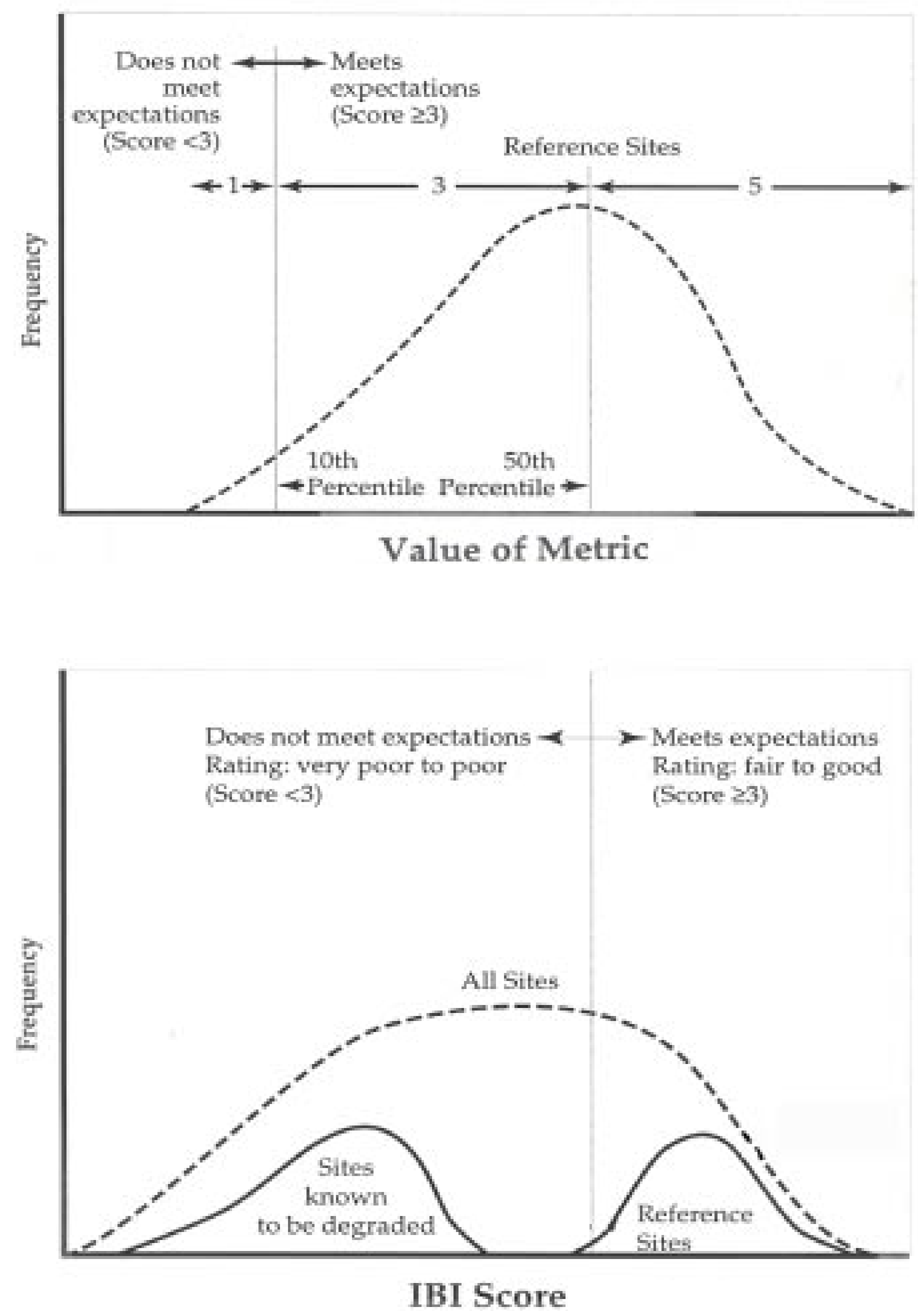


Figure 2-3. Schematic illustration of the process used to derive and interpret scores for the MBSS fish Index of Biotic Integrity (IBI). Scores are based on the distribution of reference sites, as depicted in the top figure. The bottom figure shows hypothetical reference sites in the context of other hypothetical sites, including those with known degradation.

to identify values that are outside the natural expectation for reference sites. This approach is consistent with the likelihood that in Maryland, even reference sites have some degree of anthropogenic impact.

To test the discriminatory power of each candidate metric, we evaluated the degree of overlap between metric values at reference and degraded sites by examining the number of sites scoring above and below the lower threshold. A classification efficiency was calculated as the percent of reference sites with values scoring $\geq$ 3 plus degraded sites scoring < 3, out of the total number of sites evaluated. Reference sites misclassified as degraded (score < 3) and degraded sites misclassified as reference (score $\geq$ 3) make up the remainder of the sites. A high classification efficiency indicates a small amount of overlap between values for reference and degraded sites. In addition to overall classification efficiencies, classification efficiencies were also reported separately for reference and degraded sites.

2.6 COMBINING METRICS INTO AN INDEX

To develop an overall index, different combinations of metrics were constructed and the performance of each evaluated. For each combination, an index was calculated as the mean of the metrics selected. The resulting index was scaled from 1 to 5, as were individual metrics. Classification efficiencies of different metric combinations (indices) were calculated as above.

After evaluating the performance and ecological relevance of each metric, we examined the best-performing variation of each metric as final candidates for inclusion in the index. For example, once the number of benthic species was selected based on its strong classification efficiency, the number of darter species could not be added. A core set of metrics were identified; additional metrics were added in various combinations until a single recommended combination (index) was identified for each stratum.

3. RESULTS

3.1 STRATA DETERMINATION

Three distinct geographic strata, corresponding to physiographic region and river basin boundaries, were identified via cluster analysis and MANOVA as having distinctly different naturally occurring species assemblages that corresponded with physiographic region and river basin boundaries (Figure 3-1). None of the other physical variables showed any clear correspondence with sites grouped by biological assemblage similarity. Based on these analyses, three strata were used throughout the remaining steps of the IBI development process: the Coastal Plain, Eastern Piedmont, and Highlands regions (Figure 3-2). Along with the expected Coastal Plain group, the cluster analysis showed a clear difference between sites in the Highlands (including the Middle Potomac basin and areas to the west) and Eastern Piedmont (including the Gunpowder, Patapsco, Patuxent, and other basins draining to the Chesapeake Bay). In addition, a comparison of fish abundance and species richness suggested some differences between the Eastern Piedmont and Highlands (Figure 3-3), providing additional evidence for separating these two groups (formerly grouped as "non-Coastal Plain" in the provisional IBI).

These geographic strata are coincident with aggregations of ecoregions (Omernik 1987) and the physiographic provinces (Maryland Geological Survey; Reger 1995) developed for Maryland. The Coastal Plain fish species are to a large degree distinct from those found in the higher gradient Highlands. The Eastern Piedmont appears to be a more speciose region, able to support many of the fishes found in the Highlands, but also a number of species rarely or not found in the Highlands.

The Coastal Plain includes nearly all of Maryland's eastern shore plus portions of the western shore basins below the fall line. The Eastern Piedmont stratum includes the central Maryland basins above the fall line that drain to the Chesapeake Bay and tidal Potomac. Within the Potomac Washington Metro basin, the boundary reflects the division of Potomac tributaries above Great Falls (Highlands) and below it (Eastern Piedmont), a distinction noted when cluster analysis groups were mapped. Conewago Creek, near the Pennsylvania state line, was provisionally classified in the Eastern Piedmont, because Conewago is part of the Susquehanna drainage, although reclassification of Conewago area into the Highlands region might be warranted in the future. Several sites near the fall line, previously classified as Coastal Plain according to an older map of the state's physiographic provinces, were reclassified as Piedmont, following the more recent physiographic map developed by Maryland Geological Survey (Reger 1995). The Highlands stratum includes the remainder of central and western Maryland, including the Appalachian Plateau, Valley and Ridge, Blue Ridge, and westernmost part of the Piedmont physiographic region.

The cluster analysis did not support further subdivision of these site groupings into finer geographic strata, for example, into other ecoregions or subecoregions (White 1996). In addition, although coldwater and coolwater stream systems in Maryland frequently differ in species abundance and composition from warmwater systems, a separate stratum of cold/coolwater sites was not clearly distinguishable in this analysis. Although some brook trout sites showed a slight tendency to group

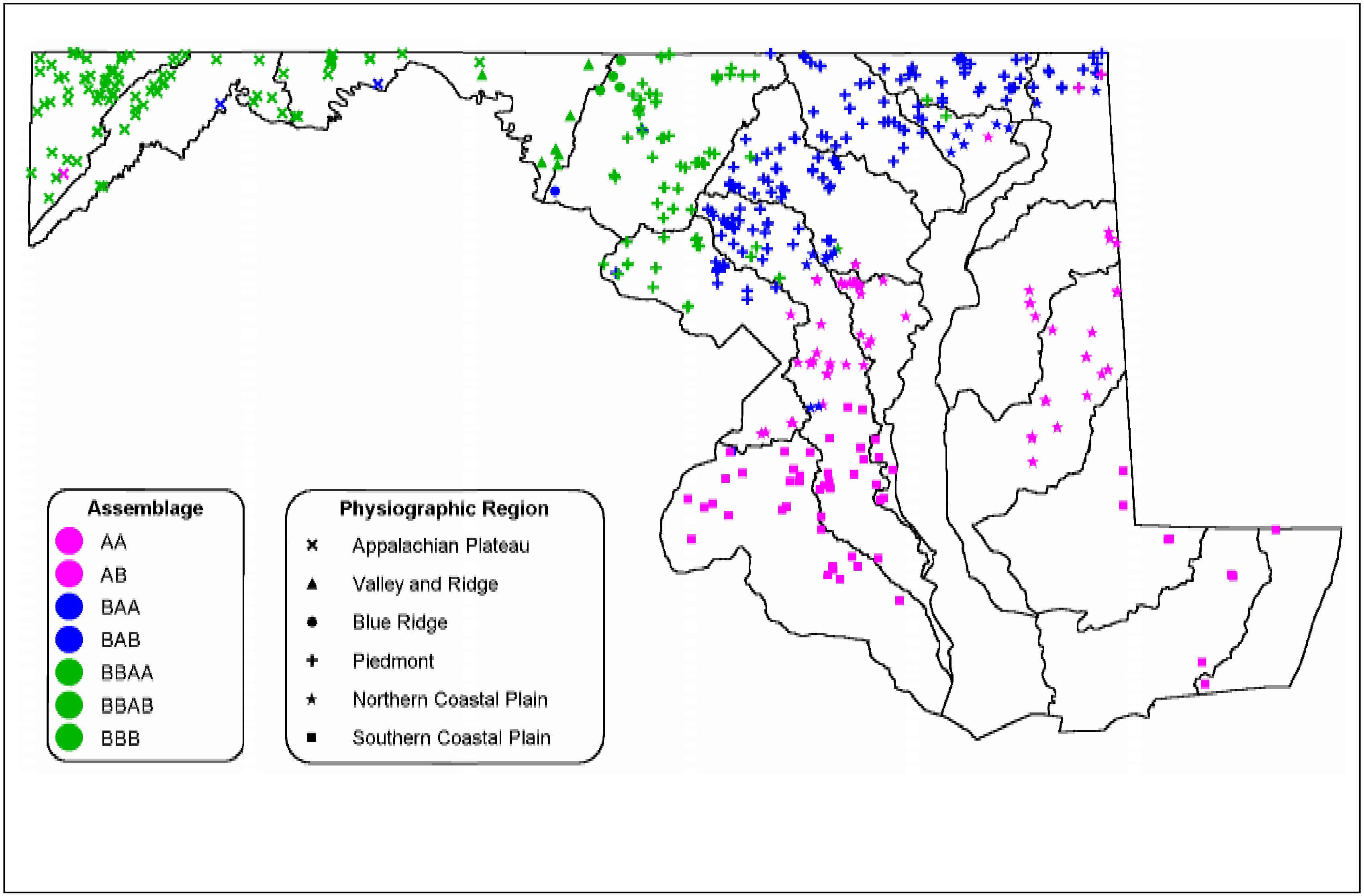


Figure 3-1. Fish assemblage groupings identified in cluster analysis (Canberra metric, n=438 sites). Three major groups (shown as red, blue, and green) corresponded with drainage basins (black lines) and physiographic region boundaries (noted by symbols).

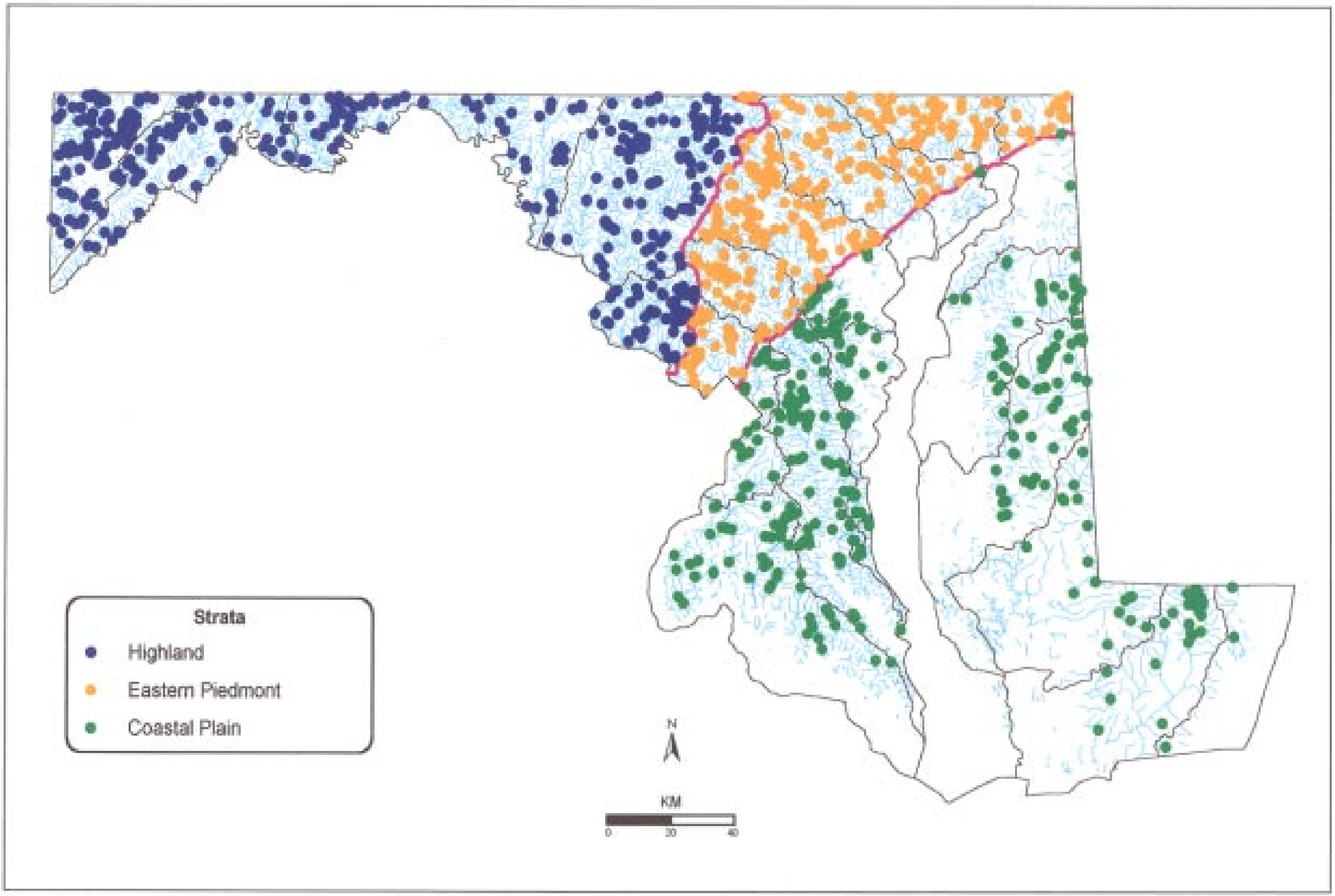


Figure 3-2. The three geographic regions used for the derivation of the fish Index of Biotic Integrity: Coastal Plain, Eastern Piedmont, and Highlands. The map shows locations of all sites in the combined IBI calibration/validation data set.

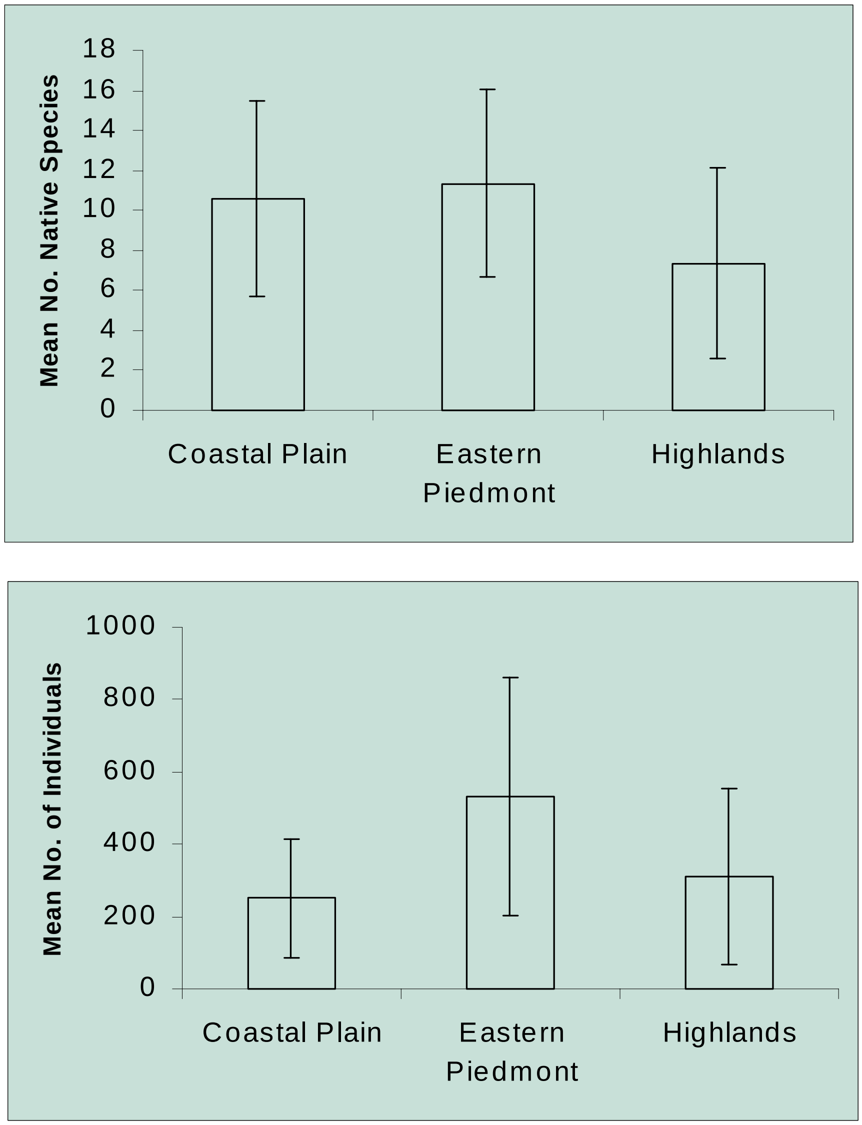


Figure 3-3. Comparison of mean species richness and fish abundance per site among the three proposed strata. Error bars indicate ± standard deviation.

together, there were no apparent physical or ecoregional covariates. There may not have been sufficient numbers of cold/cool water sites in the data set to full represent this stream type. Also, it is difficult to isolate specific physical or geographic parameters that uniquely identify coldwater streams, barring long-term temperature data (not available at present). In the future, continuous temperature data will be used to examine ways to further characterize coldwater systems and develop biological indicators specifically tailored to this habitat type. Characterization of a sufficient number of blackwater streams is also needed to develop indicators appropriate for that habitat.

3.2 METRIC EVALUATION

Using the calibration data set, the 41 candidate metrics were evaluated using statistical tests (Mann-Whitney and Kolmogorov-Smirnov), graphical analysis, and classification efficiencies. When appropriate, metrics were adjusted for watershed areas using regression relationships. Metrics were tested in each of the three strata: Coastal Plain, Eastern Piedmont, and Highlands.

Among the 41 metrics tested, many exhibited a significant difference between reference and degraded sites, passing the Mann-Whitney U and Kolmogorov-Smirnov tests ($p < 0.05$), with results varying by strata. Statistical test results along with classification efficiency (CE) values are shown in Table 3-1. A number of metrics passed both statistical tests ($P < 0.05$): 22 of the 41 metrics in the Coastal Plain, 33 in Eastern Piedmont, and 28 in Highlands. Classification efficiencies for individual metrics ranged from 45 to 73% in the Coastal Plain, 62 to 93% in Eastern Piedmont, and 64 to 90% in Highlands.

Calculation of metric scoring thresholds was based on the distribution of values at reference sites (lower threshold at 10th percentile, upper threshold at 50th percentile). In a few cases where the lower threshold equaled 0, the scoring criterion was adjusted from ≥ 0 to > 0. For example, a site with 0 intolerant species would score 1, whereas a site with one or more intolerant species would score 3. Sites with a number of intolerant species $\geq$ the upper threshold would still score 5.

The complete list of 41 metrics contained multiple variations for some metrics (e.g., number of species vs. number of native species to describe species richness). Results of statistical tests and classification efficiencies were used to select the best-performing variation of each. Table 3-2 lists these 14 best-performing metrics. From this list, only metrics passing both statistical tests were used in building metric combinations (as described in Section 3.3 below). Among metrics passing statistical tests, low threshold values were noted in a few cases:

	lower threshold	upper threshold
Percent insectivores (Eastern Piedmont)	0	0.80%
Percent top predators (Eastern Piedmont)	0	0.60%
Number of anomalies per fish examined, excluding blackspot and other visible external parasites (Highlands)	0.30	0.08

These metrics were excluded from further consideration in these strata.

Table 3-1. Evaluations of individual candidate metrics used in IBI development. Results of Mann-Whitney U and Kolmogorov-Smirnov (K-S) tests of differences between reference and degraded sites (+ indicates significant difference, $p < 0.05$). Individual classification efficiencies (CE) are the percentage of reference and degraded sites correctly classified by each metric.

Metric	Coastal Plain				Eastern Piedmont				Highlands			
	Mann-Whitney (p)	K-S (p)	Passes both tests	CE (%)	Mann-Whitney (p)	K-S (p)	Passes both tests	CE (%)	Mann-Whitney (p)	K-S (p)	Passes both tests	CE (%)
Species Richness and Composition												
Number of species*	0.00004	0.0007	+	61	<0.00001	0.0001	+	93	0.15	0.0051		76
Number of native species*	0.00007	0.0009	+	62	<0.00001	0.0001	+	86	0.11	0.0045		74
Number of darter species*	0.07	0.0168		45	<0.00001	0.0001	+	64	0.44 x	0.43		64
Number of darter and sculpin species*	0.06	0.0168		45	<0.00001	0.0001	+	91	0.0013	0.0001	+	80
Number of benthic fish species*	0.001	0.0017	+	68	<0.00001	0.0001	+	93	0.0012	0.0001	+	80
Number of darter, sculpin, and madtom species*	0.0216	0.0551		45	<0.00001	0.0001	+	93	0.0012	0.0001	+	80
Number of sunfish species*	0.0047	0.0271	+	45	0.0154	0.0112	+	64	0.85 x	0.92		64
Number of sucker species*	0.29	0.81		45	0.00006	0.0001	+	64	0.84 x	0.57		64
Percent white sucker and northern hogsucker	0.34	0.57		54	0.08	0.21		64	0.0004	0.0012	+	77
Percent round-bodied suckers	0.0378	0.19		45	0.00004	0.001	+	64	0.08	0.81		64
Indicator Species												
Number of intolerant species (literature-based)*	0.0157	0.06		66	<0.00001	0.0001	+	91	0.0028	0.0018	+	71
Number of intolerant species (based on data)*	0.0007	0.0085	+	70	<0.00001	0.0001	+	91	<0.00001	0.0001	+	83
Percent green sunfish	0.24	0.99		49	0.4	0.37		69	0.00004	0.004	+	76
Percent tolerant fish (literature-based)	0.0084	0.0193	+	62	0.00002	0.0001	+	88	<0.00001	0.0001	+	86
Percent tolerant fish (based on data)	0.0058	0.0052	+	68	0.00005	0.0001	+	88	<0.00001	0.0001	+	87
Percent white sucker	0.31	0.49		54	0.12	0.21		62	0.0003	0.0012	+	77
Percent eastern mudminnows	0.0125	0.0014	+	58	0.23	1		67	<0.00001	0.0117	+	79
Percent creek chub	0.82	0.99		46	0.4	0.17		69	0.0006	0.004	+	76
Percent pioneering species	0.0128	0.0102	+	60	0.00002	0.0001	+	91	<0.00001	0.0001	+	80
Percent abundance of dominant species	0.00004	0.005	+	61	<0.00001	0.0001	+	88	0.0043	0.0093	+	76

* values were adjusted for watershed area

Table 3-1 (Continued).

Metric	Coastal Plain				Eastern Piedmont				Highlands			
	Mann-Whitney (p)	K-S (p)	Passes both tests	CE (%)	Mann-Whitney (p)	K-S (p)	Passes both tests	CE (%)	Mann-Whitney (p)	K-S (p)	Passes both tests	CE (%)
Trophic Composition												
Percent omnivores	0.69	0.89		49	<0.00001	0.0001	+	88	<0.00001	0.0001	+	80
Percent generalists and omnivores	0.0156 x	0.0314		50	<0.00001	0.0001	+	88	<0.00001	0.0001	+	79
Percent generalists, omnivores, and invertivores	0.0082	0.0168	+	60	<0.00001	0.0001	+	88	<0.00001	0.0001	+	87
Percent omnivores and invertivores	0.00002	0.0001	+	66	0.00001	0.0001	+	83	<0.00001	0.0001	+	83
Percent insectivorous cyprinids	no value	no value		45	no value	no value		64	no value	no value		64
Percent insectivores	0.22	1		45	0.00002	0.0004	+	79	<0.00001	0.0001	+	90
Percent invertivores	0.0298 x	0.12		47	0.0014	0.0023	+	83	0.58	0.78		64
Percent insectivores and invertivores	0.0315 x	0.12		47	0.00004	0.001	+	83	<0.00001	0.0001	+	80
Percent top predators	0.08	0.19		45	0.0003	0.0019	+	76	0.19 x	0.46		64
Fish Abundance and Condition												
Number of individuals per square meter	0.0002	0.0005	+	69	0.0023	0.03	+	76	0.68 x	0.0073		74
Number of individuals*	0.0006	0.0006	+	72	0.0015	0.0044	+	79	0.42 x	0.0012		74
Number of individuals, excluding introduced species*	0.0008	0.0008	+	66	0.0023	0.0044	+	81	0.4 x	0.0012		74
Number of individuals, excluding tolerants*	0.0001	0.0009	+	62	<0.00001	0.0001	+	88	0.00003	0.0001	+	79
Biomass per square meter	<0.00001	0.0001	+	73	0.0004	0.001	+	83	0.07	0.0007		76
Biomass*	0.00001	0.0001	+	70	0.00002	0.0001	+	86	0.0265 x	0.0007		79
Number of anomalies, including blackspot and other	0.3	0.24		50	0.0492	0.06		71	0.0003	0.0051	+	73
Number of anomalies, excluding blackspot and other	0.2	0.36		49	0.0525	0.16		74	0.00004	0.0004	+	80
Reproductive Function												
Percent hybrids	0.12	1		49	0.48	1		64	0.00001	0.0073	+	77
Percent lithophilic spawners	0.0309	0.07		66	<0.00001	0.0001	+	88	<0.00001	0.0001	+	83
Percent native individuals	0.24	0.18		49	0.31	0.19		67	0.0003	0.0027	+	77
Percent native species	0.43	0.4		50	0.98	0.43		67	0.0015	0.0117	+	74

Table 3-2. Summary of metric testing and selection results. List shows results of Mann-Whitney and Kolmogorov-Smirnov tests, along with classification efficiencies for 14 metrics only (best variation of each metric type).

	Coastal Plain		Eastern Piedmont		Highlands	
Metric	Passes both tests	CE (%)	Passes both tests	CE (%)	Passes both tests	CE (%)
Number of native species*	+	62	+	86		74
Number of benthic fish species*	+	68	+	93	+	80
Number of sunfish species*	+	45	+	64		64
Number of sucker species*		45	+	64		64
Number of intolerant species (based on data)*	+	70	+	91	+	83
Percent tolerant fish (based on data)	+	68	+	88	+	87
Percent abundance of dominant species	+	61	+	88	+	76
Percent generalists, omnivores, and invertivores	+	60	+	88	+	87
Percent insectivores		45	+	79	+	90
Percent top predators		45	+	76		64
Number of individuals per square meter	+	69	+	76		74
Biomass per square meter	+	73	+	83		76
Number of anomalies, excluding blackspot and other		49		74	+	80
Percent lithophilic spawners		66	+	88	+	83

* values were adjusted for watershed area

3.3 COMBINATION OF METRICS INTO AN INDEX

Five metrics passed statistical tests in all three strata: number of benthic fish species, number of intolerant species (new metric based on DNR analysis of monitoring data), percent tolerant individuals, percent abundance of dominant species, and percent generalist-omnivore-invertivores. These five metrics were selected as a core combination and tested in each stratum, discriminating between reference and degraded sites with 69 to 88% classification efficiency (Table 3-3). Within each of the three strata, additional metrics were then added that (1) passed both statistical tests within the stratum, (2) when combined with the core metrics, improved the overall performance of the index, and (3) filled a category of metric that was missing (e.g., abundance and condition, trophic composition) where possible. Several iterations of this analysis were run. Our recommendation for the best combinations of metrics, or indices, is presented in Table 3-4.

Classification efficiencies in each of the three strata improved when additional metrics were added to the core group (Table 3-4). Eastern Piedmont had the highest overall classification efficiency (90%), followed by Highlands (86%) and Coastal Plain (74%). Among reference sites alone, classification efficiencies were 84 to 89%. This is consistent with the approach that misclassification of reference sites should be minimized (i.e., identification of degraded sites should be conservative).

The effects of adding other metrics to the common five-metric core group varied among strata, as would be expected given their ecological differences. Percent insectivores and percent lithophilic spawners boosted index performance in the Highlands, while the species richness, abundance, and biomass metrics performed poorly, perhaps reflecting the influence of smaller or cooler streams. Number of native species, abundance per square meter, and biomass per square meter improved index performance in the Coastal Plain. Nearly all metrics worked well in the Eastern Piedmont, and the final set selected for high classification efficiency–number of native species, abundance, biomass, and percent lithophilic spawners– encompasses a combination of good performers from the other two strata.

To investigate the influence of individual metrics, a “bare-bones” or “minimal” index was also evaluated. This minimal index consisted of percent tolerant individuals, number of intolerant species, and either abundance per square meter or biomass per square meter in Coastal Plain and Eastern Piedmont; the same combinations without abundance or biomass were tested in Highlands. These 3-metric and 2-metric combinations performed well (CE 74 to 93%; Table 3-5), indicating that information on the composition of tolerant/intolerant species and, in some cases, abundance or biomass may be driving IBI scores. As we found before (Roth et al. 1998a), the discriminatory power of the IBI appears to be derived from only a small number of metrics. Additional metrics that are used here, and are traditionally included in IBIs, can add to the robustness of the index (applicability beyond the development data set) and provide a broader ecological basis for interpretation.

The final formulation of the IBI for each of the three strata, with scoring thresholds, is given in Table 3-6.

Table 3-3. Classification efficiencies for the core combination of metrics, using five metrics common to all strata.

	Coastal Plain	Eastern Piedmont	Highlands
	Core	Core	Core
Overall classification efficiency	69	88	84
Classification efficiency at reference sites	76	82	87
Classification efficiency at degraded sites	63	100	80
Species Richness and Composition			
Number of benthic fish species*	x	x	x
Indicator Species			
Number of intolerant species (based on data) *	x	x	x
Percent tolerant fish (based on data)	x	x	x
Percent abundance of dominant species	x	x	x
Trophic Composition			
Percent generalists, omnivores, and invertivores	x	x	x

* values were adjusted for watershed area

Table 3-4. Classification efficiencies for metric combinations (indices), using final formulation of five core metrics common to all strata, plus additional metrics specific to individual strata.

	Coastal Plain	Eastern Piedmont	Highlands
	Final	Final	Final
Overall classification efficiency	74	90	86
Classification efficiency at reference sites	85	89	84
Classification efficiency at degraded sites	66	93	88
Species Richness and Composition			
Number of native species*	x	x	
Number of benthic fish species*	x	x	x
Indicator Species			
Number of intolerant species (based on data) *	x	x	x
Percent tolerant fish (based on data)	x	x	x
Percent abundance of dominant species	x	x	x
Trophic Composition			
Percent generalists, omnivores, and invertivores	x	x	x
Percent insectivores			x
Fish Abundance and Condition			
Number of individuals per square meter	x	x	
Biomass per square meter	x	x	
Reproductive Function			
Percent lithophilic spawners		x	x

* values were adjusted for watershed area

Table 3-5. Classification efficiencies for 2- and 3-metric combinations ("minimal" indices).

	Coastal Plain		Eastern Piedmont		Highlands
	3-metric		3-metric		2-metric
Overall classification efficiency	76	74	93	90	84
Classification efficiency at reference sites	88	82	96	89	87
Classification efficiency at degraded sites	66	68	87	93	80
Indicator Species					
Number of intolerant species (based on data) *	x	x	x	x	x
Percent tolerant fish (based on data)	x	x	x	x	x
Fish Abundance and Condition					
Number of individuals per square meter	x		x		
Biomass per square meter		x		x	

* values were adjusted for watershed area

Table 3-6. Metrics and scoring criteria for the recommended final fish IBI. Some metrics[a] were adjusted for watershed area, based on linear relationships[b] between the metric and log(watershed area) in acres

	Scoring criteria 5	3	1
Coastal Plain			
Number of native species[a]	Criteria vary with stream size (see below)		
Number of benthic fish species[a]	Criteria vary with stream size (see below)		
Number of intolerant species[a]	Criteria vary with stream size (see below)		
Percent tolerant fish	≤ 50	50 < x ≤ 93	> 93
Percent abundance of dominant species	≤ 33	33 < x ≤ 78	> 78
Percent generalists, omnivores, and invertivores	≤ 92	92 < x < 100	100
Number of individuals per square meter	≥ 0.79	0.42 ≤ x < 0.79	< 0.42
Biomass (g) per square meter	≥ 9.9	3.6 ≤ x < 9.9	< 3.6
Eastern Piedmont			
Number of native species[a]	Criteria vary with stream size (see below)		
Number of benthic fish species[a]	Criteria vary with stream size (see below)		
Number of intolerant species[a]	Criteria vary with stream size (see below)		
Percent tolerant fish	≤ 41	41 < x ≤ 65	> 65
Percent abundance of dominant species	≤ 30	30 < x ≤ 52	> 52
Percent generalists, omnivores, and invertivores	≤ 86	86 < x ≤ 99.7	> 99.7
Number of individuals per square meter	≥ 0.81	0.35 ≤ x < 0.81	< 0.35
Biomass per square meter	≥ 8.0	3.7 ≤ x < 8.0	< 3.7
Percent lithophilic spawners	≥ 62	22 ≤ x < 62	< 22
Highlands			
Number of benthic fish species[a]	Criteria vary with stream size (see below)		
Number of intolerant species[a]	Criteria vary with stream size (see below)		
Percent tolerant fish	≤ 28	28 < x ≤ 71	> 71
Percent abundance of dominant species	≤ 49	49 < x ≤ 91	> 91
Percent generalists, omnivores, and invertivores	≤ 49	49 < x ≤ 92	> 92
Percent insectivores	≥ 48	8 ≤ x < 48	< 8
Percent lithophilic spawners	≥ 70	42 ≤ x < 70	< 42

Table 3-6. Continued

[(a)] Adjusted value = observed value/expected value, where expected value = m * log(watershed area in acres) + b.

	Scoring criteria		
	5	3	1
Coastal Plain			
Number of native species - Adjusted value	≥ 1.06	0.53 ≤ x <1.06	< 0.53
Number of benthic fish species - Adjusted value	≥ 1.06	0 < x < 1.06	0
Number of intolerant species - Adjusted value	≥ 0.34	0 < x < 0.34	0
Eastern Piedmont			
Number of native species - Adjusted value	≥ 1.02	0.56 ≤ x < 1.02	< 0.56
Number of benthic fish species - Adjusted value	≥ 0.99	0.50 ≤ x < 0.99	< 0.50
Number of intolerant species - Adjusted value	≥ 0.59	0.18 ≤ x < 0.59	< 0.18
Highlands			
Number of benthic fish species - Adjusted value	≥ 1.03	0.33 ≤ x < 1.03	< 0.33
Number of intolerant species - Adjusted value	≥ 0.73	0.23 ≤ x < 0.73	< 0.23

[(b)] Slope and intercept values for selected metrics, based on linear regression relationships between metric and log(watershed area) in acres

	slope (m)	intercept(b)
Coastal Plain		
Number of native species	6.5936	-13.0055
Number of benthic fish species	1.5743	-3.3929
Number of intolerant species	2.1485	-5.286
Eastern Piedmont		
Number of native species	5.5701	-8.1135
Number of benthic fish species	1.3245	-2.6437
Number of intolerant species	4.4052	-8.8991
Highlands		
Number of benthic fish species	1.6067	-3.5202
Number of intolerant species	3.0723	-7.3029

3.4 VALIDATION OF IBI USING INDEPENDENT DATA

After defining reference and degraded sites, the 1994-1997 data set had been subdivided by randomly assigning 2/3 of the sites as "calibration" sites and reserving 1/3 for independent validation of the IBI. Using this validation data set (Table 2-3), the classification efficiency of the final IBI was evaluated. Validation data were tested using the same steps employed in calibration: (1) identifying reference and degraded sites, (2) scoring of metrics using the thresholds in Table 3-6, and (3) evaluating final indicator performance.

The final index validated well, correctly classifying 94% of sites in the Eastern Piedmont, 75% of sites in the Highlands, and 72% of sites in the Coastal Plain. Complete validation results are presented in Table 3-7. Because the classification efficiencies among validation sites in the Highlands and Coastal Plain were slightly lower than at calibration sites, lists of physical habitat, chemical, and land use data for the calibration and validation data sets were examined to determine what factors might be influencing index performance.

Among validation sites in the Highlands, no degraded sites were misclassified. Most of the misclassified reference sites in this region were small, shallow streams (average thalweg depth < 20 cm). Only one of these sites had brook trout. When shallow streams were excluded, overall classification efficiency in Highlands validation improved from 75 to 93%. This suggests that IBI results reported for shallow Highlands streams should be examined carefully, since low IBI scores may be sensitive to small stream size, not merely to degraded conditions.

Among Coastal Plain validation sites, degraded sites were more likely to be misclassified than reference sites, often because of higher than expected species richness at degraded sites. Individual metrics had poorer performance (lower CE values) in this stratum than others, and graphical analysis showed greater overlap in metric values at reference and degraded sites. This result is not unexpected, given the high degree of human impact in the Coastal Plain. One possible reason for the high misclassification rate is that essentially all Coastal Plain systems have been greatly modified and thus bear less resemblance to their natural condition than do streams elsewhere in the state.

3.5 RECOMMENDED FORMULATION FOR FISH IBI

The recommended final fish IBIs for the three strata are listed in Table 3-6 along with thresholds for metric scoring. Additional information about the metrics is given in Table 3-8. These metrics and values were subsequently adopted for the MBSS statewide stream assessment (Roth et al. 1999) and are currently in use by Maryland DNR.

Table 3-7. Classification efficiencies of final IBI in each stratum using the validation data set

	Coastal Plain	Eastern Piedmont	Highlands
	Final	Final	Final
Overall classification efficiency	72	94	75
Classification efficiency at reference sites	90	91	68
Classification efficiency at degraded sites	64	100	100

Table 3-8. Description of fish IBI metrics
Number of native species (adjusted for watershed area) **-** Total number of native fish species; adjusted for watershed area (see Table 3-6b). Fishes were classified as native or introduced to Chesapeake Bay or Youghiogheny/Ohio River drainage.
Number of benthic fish species (adjusted for watershed area) **-** The number of fish species that reside primarily on the stream bottom, adjusted for watershed area (see Table 3-6b). Benthic fishes include all darters (*Etheostoma* spp., *Perca* spp.), sculpins (*Cottus* spp.), madtoms (*Noturus* spp.), and lampreys (*Petromyzon* spp., *Lampetra* spp.).
Number of intolerant species (adjusted for watershed area) **-** The number of fish species rated as intolerant of anthropogenic stress, adjusted for watershed area. Tolerance ratings (intolerant, tolerant) were based on statewide analysis comparing species occurrences with presence/absence of anthropogenic stressors.
Percent tolerant fish **-** The percentage of individuals rated as tolerant to anthropogenic stress.
Percent abundance of dominant species **-** The percentage of individuals within the single most abundant (dominant) species at a site.
Percent generalists, omnivores, and invertivores **-** The percentage of individuals classified into the trophic groups of generalist, omnivore, or invertivore; these are the most general of all feeding habits. Invertivores eat insects and other invertebrates including crustaceans, mollusks, and worms. Omnivores consume two or more food types (insects, invertebrates other than insects, fish, plankton, algae, vascular plants, and detritus) with the exception of the combination of invertebrates and fishes. Generalists eat both invertebrates and fishes but not other food items.
Percent insectivores **-** The percentage of individuals classified into the group insectivore; this is a specialized trophic group, feeding almost exclusively on insects.
Number of individuals per square meter **-** The number of individuals captured at a site, divided by the surface area fished. Surface area was computed as length of stream fished (usually 75 m) multiplied by average stream width.
Biomass (g) per square meter **-** Total mass in grams of fish captured at a site, divided by the surface area fished.
Percent lithophilic spawners **-** The percentage of individuals reported to use rock substrates for spawning.

The metrics used in the IBI were selected because they represent attributes of the fish assemblage indicative of ecological quality. The following paragraphs discuss the ecological basis for the using these metrics.

Number of native species. The concept of species richness has been used extensively to assess the quality of ecological systems. In most cases, the number of fish species supported by streams of a given size in a given region decreases with environmental degradation (Karr et al. 1986). The reduction in number of species may result from the reduced diversity of habitats or the loss of species that are sensitive to pollutants or other human-induced impacts. Introduced species (Table 2-2) are not included in this metric because the presence of these species may result in a higher species number than would naturally be found in a given stream. In addition, the species richness value for a site where species have been introduced would not reflect the lowered richness that may result from human disturbance at the site. Leidy and Fiedler (1985) found that species richness increased at sites with moderate human disturbance mostly as a result of introduced species. There are some potential exceptions to this rule. For example, minimally disturbed coldwater systems, dominated by salmonids and sculpin, tend to have naturally low numbers of species; increases in temperature may increase the total number of species as native, warmwater species invade.

Number of benthic fish species. Benthic fish species are sensitive to degradation of stream benthic habitats because they have specific requirements for reproducing and feeding on the stream bottom (Page 1983). Benthic habitats can be degraded by channelization, siltation, or the reduction of dissolved oxygen and are often degraded in streams with watersheds that contain large amounts of impervious surface. Berkman and Rabeni (1987) documented reduced abundance of benthic insectivores in streams with increased amounts of silt in riffles. Darter, sculpin, madtom, and lamprey species were included as benthic specialists in this metric.

Number of intolerant species and **percent tolerant individuals.** By definition, intolerant species are among the first to be affected by perturbations (Jenkins and Burkhead 1993, Pflieger 1975, Smith 1979, Trautman 1981). As specific habitats required by habitat specialists are degraded, the relative abundance of tolerant habitat generalists becomes greater.

Percent abundance of the dominant species. The percentage of individuals belonging to the dominant (tolerant) taxa in the fish community is likely to increase as the amount and extent of degradation increases. As intolerant species become less abundant in degraded streams, tolerant species increase in relative abundance and may become the dominant taxa (Karr et al. 1986). This metric was calculated as the percent contribution of the single dominant fish species to the total number of individuals at a site.

Percent of individuals as generalists, omnivores, or invertivores. The dominance of generalist feeders increases as specific food sources become less reliable, i.e., when degraded conditions reduce the abundance of particular prey items. An opportunistic foraging strategy makes generalists more successful than specialized foragers, because they are better suited to a shifting food base (common with degraded conditions) than are more specialized feeders (Karr et al. 1986).

Percent of individuals as insectivores. This metric takes into account the response of fishes to impacts on lower trophic levels. The fact that fewer insectivorous fishes are collected in degraded streams probably reflects decreases in the supply of preferred insects that result from pollution or degraded habitat quality (Karr et al. 1986).

Abundance (number of individuals) per square meter. Degraded streams are generally expected to yield fewer individuals than less severely impacted streams. In streams of similar size, those with greater heterogeneity of habitat generally contain larger numbers of individuals than streams where anthropogenic impact has resulted in more homogeneous habitat. In addition, streams with degraded chemical or habitat quality that support only tolerant species are likely to have lower overall numbers of fishes. One notable exception is the elevated abundance of fishes, particularly of tolerant species, in the presence of excess nutrients.

Biomass per square meter. The biomass that a stream can accommodate is a function of the quantity and quality of available stream habitat. As with abundance, the biomass in a stream is likely to be lower in degraded streams than in high quality streams. In general, more and larger fishes are expected in higher quality streams. The presence of larger individuals of a species likely indicates that the stream has had a history of good stream quality. Although aggregate fish biomass is included in the Coastal Plain IBI, note that this metric could be influenced by the presence of a few tolerant individuals of large body size, such as common carp (*Cyprinus carpio*), in which case biomass may not accurately reflect high biotic integrity. This influence was not found in our dataset, but it could be important in other cases. An alternative metric might only include biomass for particular species.

Percent of individuals as lithophilic spawners. Lithophilic spawners (Balon 1975) utilize rocks, rubble, or gravel substrates for egg deposition. Because they require clean spawning substrates and may use interstitial spaces, lithophils are particularly susceptible to siltation. Since silt is likely the most common stream pollutant in the state of Maryland, this metric may be useful in identifying streams that are degraded with substantial silt loads.

4. DISCUSSION

4.1 IMPROVEMENTS IN AND LESSONS FROM IBI DEVELOPMENT

The final IBI performed well in both the calibration and validation analyses. Overall, calibration sites were correctly classified in 82% of cases. Classification efficiencies within strata ranged from 74 to 86% for calibration sites and 72 to 94% in independent validation. Consistent with our intent to avoid giving low scores to good sites, classification efficiencies were particularly high among reference sites.

Although much of the structure of the provisional IBI was retained in the final refined IBI, several improvements were made through new analyses. Additional quantitative analysis compared new metrics and new metric combinations to ensure that the final indices would most effectively classify reference and degraded sites. Analyses of data on individual species' tolerances to degradation provided a stronger numerical basis for two metrics — the relative proportion of tolerant individuals and the number of intolerant species.

Most importantly, we used the statewide data for the first time to characterize geographic differences, adding a third stratum to better capture the range of variability in Maryland streams. While the fine-tuning of final metric selection for each of the three geographic regions was based on their performance in the statewide data set, it was consistent with ecological professional judgment about the regions. For example, species richness and fish abundance are not included in the Highlands IBI, and percent lithophilic spawners is included only in the Eastern Piedmont and Highlands, not in the Coastal Plain, where silt and sand substrates naturally predominate.

Applying physiochemical and land use criteria proved to be an effective means of objectively identifying reference and degraded sites by characterizing a broad range of expected impacts including urban land use. Reference sites were selected to represent regional natural habitats, also referred to as "minimally impacted" conditions. We recognize that virtually no streams in Maryland are entirely undisturbed by human activities. Atmospheric deposition of contaminants alone reaches all parts of the State, few streams have natural temperature regimes, and more than 1,000 man-made barriers to fish migration have been documented in Maryland. Therefore, our reference conditions should not be viewed as completely natural or pristine. They are, however, a representative sample of the best streams that currently exist in the State. Whether these conditions are the best attainable depends on future restoration activities and the goals of DNR and the public.

While some have suggested that modified reference conditions be developed for situations where human impact is pervasive and unlikely to be reversed (such as urban streams), we have not taken this approach. Instead, our reference sites were selected to establish expectations for natural streams within each geographic region, and urban streams are rated on the same scale as other sites in the region. Although some urban streams may never recover to a level comparable to natural streams, we believe our references conditions are the appropriate benchmark for assessing stream conditions. Appropriate management actions could be set using an intermediate IBI value as the desired goal. This strategy could facilitate restoration of heavily impacted streams to a stream condition that is practical and attainable (given their history of degradation and current level of watershed development), while still indicating the degree to which this restoration falls short of natural conditions.

Our quantitative analyses also helped evaluate why the IBI performed as it did. Specifically, we found that only a few key metrics need to be included in the IBI to achieve a high degree of discriminatory power. Our approach differs from many other IBI development methods in that we started with a small number of metrics and added other metrics only if they did not sacrifice performance of the overall IBI. Many IBI applications have included a standard set of ten to twelve metrics (with minor regional modifications or substitutions) without testing to determine if the metrics help or hinder discriminatory ability. In our case, even a two- or three-metric IBI (characterizing fish abundance and predominance of tolerant species) was able to effectively distinguish site quality. Nonetheless, we recognize that the addition of more metrics probably improves IBI robustness by enhancing discrimination in rare cases or in addressing situations that may be encountered in future applications of the index. Angermeier and Karr (1986) also noted wide variation in the contributions of individual metrics to IBI performance, with the relative importance of particular metrics differing among regions and degradation gradients. We suspect that IBIs developed by other investigators also derive most of their discriminatory power from a relatively small subset of their metrics. While we acknowledge the value of retaining a sufficient number of metrics to address a variety of impacts, we also encourage careful selection of metrics based on evaluations of their individual and combined discriminatory power.

Many stream monitoring efforts have avoided sampling the smallest streams included in the probability-based MBSS design. Of the 1098 sites used for IBI development, 32% drained catchments smaller than 1,000 acres (405 hectares), while about 87% drained areas less than 12,000 acres (4,856 hectares). In contrast, the majority of streams sampled by Ohio EPA's rotating basin monitoring program have catchments larger than 12,000 acres. Ohio also incorporates special provisions for small streams, recognizing both the fundamentally different fauna and utility of assessments in small streams. For example, Ohio routinely corrects metrics for sites with less than 200 fish per 300 m (Ohio EPA 1987). We recognized that developing IBIs for these small, less species-rich streams might be difficult, but we did not want to dismiss this widespread but often overlooked resource. Nonetheless, intuition argues for a lower limit on the number of fish and species that must be sampled at a site to produce a useful IBI. Indeed, variability in IBI scores has been shown to increase with low total abundance (<400 individuals per sample), raising concerns about applying the index in small streams (Fore et al. 1994). Based on our data, we designated a minimum catchment size requirement of 300 acres (121 hectares). Future MBSS efforts may generate alternative methods for assessing these small headwaters. Assessments of small streams will continue to be important, particularly as MBSS moves to using a 1:100,000-scale base map, rather than the current 1:250,000 scale.

Although included in other IBIs, the number of anomalies was not found to be effective in discriminating between reference and degraded conditions, primarily because anomalies occurred relatively infrequently in the data set. Ohio EPA (1987) includes in their IBI the proportion of individuals with deformities, eroded fins, lesions, and tumors ("DELT anomalies"), but reports that this metric is most sensitive at highly degraded sites subject to point source impacts. At MBSS sites, the number of anomalies was generally low, probably because samples were taken from relatively small streams. Potentially, a survey of fourth-order and larger streams might involve more point source impacts and could profit from including anomalies as a metric. Although we did not include the number of anomalies as a metric in the IBI, we recommend it be reported separately as an additional indicator of fish assemblage health.

Given the natural variability of fish assemblages, even the best index is unlikely to provide a 100% classification efficiency. Ideally, rates of misclassifying both reference and degraded sites should be minimized. However, even a very useful index will likely have some overlap between reference and degraded conditions resulting from natural variability in fish assemblage composition. The best way to balance these two kinds of "misclassifications" depends on the purpose of assessment. Programs to screen sites for potential problems should produce a low misclassification of degraded sites, while efforts to identify specific sites for remedial actions should avoid mislabeling high quality sites. The Maryland IBI, which may be used to target future restoration efforts, employs the second strategy, i.e., choosing to rarely misclassify reference, or nondegraded, sites.

4.2 INTERPRETING THE IBI

Scores for the fish IBI are determined by comparing the sampled fish assemblage at each site to those found at minimally impacted reference sites, using the metrics and thresholds listed in Table 3-6. In calculating the IBI, the proper regional formulation should be used for each of the three distinct geographic areas: Coastal Plain, Eastern Piedmont, and Highlands (Figure 3-2). Individual metrics for the IBI are scored 1, 3, or 5, based on comparison with the distribution of metric values at reference sites. A score of 3 or greater is considered comparable to reference site conditions, while scores falling below this threshold differ significantly from the reference conditions, as shown in Figure 2-3. Scores for the IBI are calculated as the mean of the individual metric scores and therefore also range from 1 to 5. Some other programs have used a similar approach (e.g., Weisberg et al. 1997), while others have computed the IBI as the total of individual metric scores. For example, Karr et al. (1986) calculated IBI as the sum of 12 metric scores, with totals ranging from 12 to 60 points.

For management purposes, site-specific IBI results have been used to estimate the statewide extent of non-tidal streams in good, fair, poor, and very poor condition with respect to the biotic integrity of the fish community. Table 4-1 contains narrative descriptions for each of the IBI categories developed for the MBSS. Originally, the IBI was designed to distinguish degraded from nondegraded conditions, using the value of ≥ 3 as the threshold to characterize a site as nondegraded (i.e., having biological attributes comparable to reference sites). The additional distinctions of good vs. fair, and poor vs. very poor were developed to serve the management need for finer classifications. The highest scores (IBI 4 to 5) were designated as good (rather than excellent) recognizing that available reference sites may not represent the highest attainable condition. The assignment of scores to narrative categories is a useful method for translating scores into a form that is easily communicated. Similar approaches have been used in other IBI applications (e.g., Karr 1991, Ohio EPA 1987, Ranasinghe et al. 1996).

Table 4-1. Narrative descriptions of stream biological integrity associated with each of the IBI categories		
Good	IBI score 4.0 - 5.0	Comparable to reference streams considered to be minimally impacted. On average, biological metrics fall within the upper 50% of reference site conditions.
Fair	IBI score 3.0 - 3.9	Comparable to reference conditions, but some aspects of biological integrity may not resemble the qualities of these minimally impacted streams. On average, biological metrics are within the lower portion of the range of reference sites (10th to 50th percentile).
Poor	IBI score 2.0 - 2.9	Significant deviation from reference conditions, with many aspects of biological integrity not resembling the qualities of minimally impacted streams, indicating some degradation. On average, biological metrics fall below the 10th percentile of reference site values.
Very Poor	IBI score 1.0 - 1.9	Strong deviation from reference conditions, with most aspects of biological integrity not resembling the qualities of minimally impacted streams, indicating severe degradation. On average, biological metrics fall below the 10th percentile of reference site values; most or all metrics are below this level.

4.2.1 Special Considerations in Interpreting IBI Scores

Several basins in Maryland contain streams that can be classified as coldwater stream systems. Generally, high-quality coldwater streams are dominated by brook trout and have lower overall species richness than warmwater systems of the same area. In other parts of North America, fish IBI frameworks for coldwater and coolwater streams have been tailored to account for their unique biological characteristics (Leonard and Orth 1986, Lyons et al. 1996, Mundahl and Simon 1999). In Maryland, three regional fish IBIs were used to assess MBSS sites. However, because the IBI may underrate coldwater streams owing to their naturally low species diversity, the presence of brook trout should be used as a secondary indicator in interpreting fish IBI scores. In the statewide stream assessment (Roth et al. 1999), sites where brook trout were present and fish IBI scores were less than 3 were excluded from analysis and reported as "not rated." This situation was rare (14 sites), constituting only 1.5% of all sites and 20% of brook trout sites.

Other types of natural variability should be considered when applying the IBI, especially in areas expected to differ in species richness and diversity. Naturally acidic blackwater streams may have lower species richness and may be dominated by a few acid-tolerant species. A total of 24 MBSS sites were identified as blackwater streams, defined operationally as sites with either pH < 5 or ANC < 200 µeq/l and DOC > 8 mg/l. To avoid possibly underrating blackwater streams, the nine blackwater streams with fish IBI scores less than 3 were excluded from statewide analysis and were therefore included in the"not rated" category. Maryland DNR is considering developing separate IBIs for coldwater and blackwater stream types in the future.

Other factors that may affect fish IBI scores should be considered in interpreting scores for individual sites. Small streams with shallow stream channels may naturally support few species. Flood or drought conditions can temporarily reduce the abundance and diversity of fish in an affected stream. Dams and other barriers to fish migration can block access to formerly inhabited upstream areas. In contrast, proximity of a site to a lake, pond, swamp, or impoundment in a watershed can make a site more accessible to lentic species not typically found in the small streams sampled by the MBSS. Nearness to a large river confluence can similarly alter the pool of available species. High species richness resulting from the presence of both Coastal Plain and Piedmont species at sites along the Fall Line may result in artificially high IBI scores in this transitional area. In the final analysis, a practical IBI cannot accommodate the full range of natural variability (because an enormous number of strata would be needed), so rare stream types and unique situations need to be addressed on a case-by-case basis.

4.3 APPLICATIONS OF THE IBI

This fish IBI has been successfully used in conjunction with MBSS physical and chemical data to answer critical questions about the condition of Maryland streams and the relative impacts of human-induced stresses on the state's aquatic systems. As reported in a statewide assessment of stream conditions (Roth et al. 1999), fish IBI scores for 1995-1997 MBSS sites spanned a wide range of biological conditions, from good to very poor. A estimated 45% of stream miles statewide fell into the range of good to fair, while about 28% of stream miles showed some level of degradation (27% were not rated). IBI assessments were used to evaluate biological impacts of acidification, habitat degradation, and watershed land use.

Currently, the fish IBI is being used in conjunction with the benthic IBI to assess Maryland streams and report their status under Section 305(b) of the Clean Water Act. Protocols for applying both benthic and fish IBI results in state decisions to list degraded waters under Section 303(d) are under development, as the State considers how to best implement biological criteria to supplement existing water quality standards. Currently, the fish IBI can be used in conjunction with the benthic IBI as a screening tool for identifying potentially degraded waters. Final listing decisions will likely require collection and analysis of additional stressor data to identify likely causes of degradation needed for TMDL development.

The fish IBI has applications in resource management beyond support of regulations. IBI scores at MBSS or other sites are particularly useful important for characterizing local watersheds and can help managers identify both sites in need of restoration and high quality sites as candidates for preservation. The fish IBI has already been used as one component of a statewide biodiversity study by identifying locations with high-integrity aquatic communities (Southerland et al. 1999). Increasingly, county and local governments are using bioassessment tools such as the fish IBI in their own water quality monitoring to support stormwater management, land use planning, and other water resource management initiatives.

4.4 FURTHER EVALUATIONS OF THE IBI

As Maryland has begun to incorporate the MBSS fish IBI into the State's biological criteria, several additional analyses have been completed. Summarized below, these investigations supplement the IBI development and validation by assessing the variability in IBI scores, distributions of IBI scores at reference and degraded sites, and effects of adjusting certain reference site criteria.

4.4.1 IBI Variability

One important consideration in using IBI scores as regulatory biocriteria is the expected variability in scores. Setting appropriate thresholds for designating watersheds as impaired requires knowledge of the expected uncertainty around IBI values. Where sufficient data were available within a given area of interest (n $\geq$ 10 sites), standard statistical estimates (such as the standard deviation) were calculated from sample data and used to estimate confidence intervals for watersheds sampled in the 1995-1997 MBSS. In contrast, the variance around an individual IBI value at a site is unknown, but can be estimated from other data. To estimate variability, fish IBI data from stream reaches with two or more sites (sites within 1.0 km of one another and with similar land use, water chemistry, and physical habitat) were analyzed. Results were used to estimate the average within-segment coefficient of variation (*cv*) for fish IBI scores. Average *cv* was 0.08 (8%) of the IBI score. Additional analyses of variability at multiple scales showed that, as expected, IBI scores were more variable at larger watershed scales. Full details of this analysis are reported in Roth et al. (2000).

4.4.2 Distributions of IBI scores at reference and degraded sites

Currently the MBSS defines the IBI score as , the average of individual metric scores. Individual metrics scores are based on comparison with the distribution of metric values at reference sites, within each stratum (Coastal Plain, Eastern Piedmont, Highlands). Metrics are scored 1 (if < 10th percentile of reference value), 3 (10th to 50th percentile), or 5 ($\geq$ 50th percentile). The final IBI scores are calculated as the average of these scores. An IBI $\geq$ 3 (fair to good) indicates the presence of a biological community with attributes (metric values) comparable to those of reference sites, while an IBI < 3 (very poor to poor) means that, on average, metric values fall short of reference expectations.

This approach differs from some other IBIs, which instead use the distribution of final IBI scores at reference sites (rather than individual metric values) to establish thresholds. Under that system, a certain percentile of IBI scores at reference sites is established as the threshold for identifying sites as sites degraded. To evaluate how the scoring and interpretation of the MBSS IBI would change under this alternative approach, the 10th, 25th, and 50th percentiles of reference site IBI scores were calculated for all sites combined and by stratum using the calibration data set (Table 4-2). To quantify the current threshold (IBI =3) in terms of reference site distribution, the percentile value at IBI= 3 was also calculated. Among all reference sites (n=145), the 10th percentile was 2.46 and 50th percentile was 3.86. IBI = 3 represented the 17th percentile. In all strata, the 10th percentile was less than 3.0, and the 25th and 50th percentile were greater than 3.0 (i.e., currently the 17^{th} percentile for all reference sites).

If the 10th percentile statewide were selected as a threshold for detecting degradation, this would mean only those sites scoring < 2.46 would be considered degraded. This would represent a lowering of the current standard. We recommend instead that to be more conservative (i.e., more protective of stream condition), the MBSS IBI retain the current threshold of 3.0. In all cases, this is at least as protective, and most of the time even more protective, than the 10th percentile.

Table 4-2. Percentiles of IBI scores at reference sites, MBSS 1994-1997 fish IBI development data set

	n	**10th percentile**	**25th percentile**	**50th percentile**	**Percentile at which IBI=3**
All Reference Sites	145	2.46	3.25	3.86	17th percentile
Reference Calibration Sites - Coastal Plain	33	2.15	3.25	3.75	16th percentile
Reference Calibration Sites - Eastern Piedmont	27	2.82	3.44	3.89	12th percentile
Reference Calibration Sites - Highlands	45	2.43	3.57	3.86	16th percentile

We also evaluated the effect of changing the threshold on IBI classification efficiency. Frequency distributions of IBI scores at reference and degraded sites (Figure 4-1) were examined and classification efficiencies calculated assuming thresholds of 3.0, 10th percentile, and 25th percentile, using percentile values specific to each stratum (Table 4-3). Overall classification efficiencies were not improved using the alternative approach, except for a slight increase in the Coastal Plain using the 25th percentile (change from 74% to 77%). Generally, when the threshold was shifted down to the 10th percentile, more reference sites were correctly classified, but more degraded sites were incorrectly classified as fair-to-good. The opposite occurred when the 25th percentile of scores was used: more degraded sites were correctly classified, but more reference sites were incorrectly classified as poor-to-very-poor. The most dramatic difference was the extremely poor classification (44%) of Coastal Plain degraded sites using the 10th percentile threshold.

Based on these results, no changes to the current threshold scoring are recommended at this time. However, in future refinements of the IBI, alternative scoring approaches should be considered, with the objective of improving classification rates.

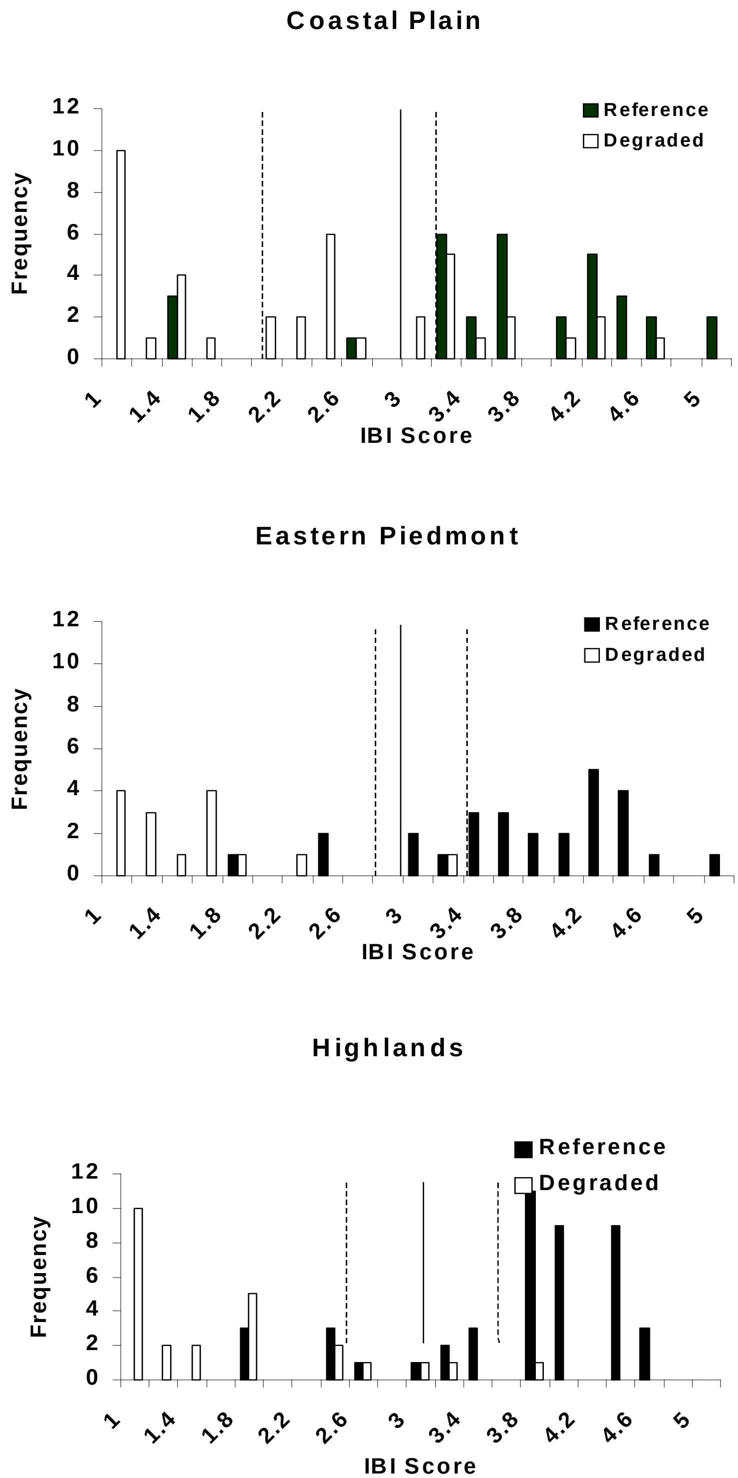


Figure 4-1. Distributions of IBI scores at reference and degraded sites in each stratum, MBSS fish IBI calibration data set. The solid vertical line indicates the threshold of 3.0; dotted lines represent the 10th (left) and 25th (right) percentile of reference site scores within the stratum.

Table 4-3. Classification efficiencies (% of calibration sites correctly classified) that would result from using different thresholds

	Classification Efficiencies Under Different Thresholds		
	3.0	**10th percentile**	**25th percentile**
Coastal Plain	**74**	**64**	**77**
Reference Sites	85	88	85
Degraded Sites	66	44	71
Eastern Piedmont	**90**	**90**	**86**
Reference Sites	89	89	78
Degraded Sites	93	93	100
Highlands	**86**	**86**	**84**
Reference Sites	84	87	78
Degraded Sites	88	84	96

4.4.3 Effects of adjusting reference site criteria

As described in Section 2.2, reference sites were chosen on the basis of objective, nonbiological factors. Evaluation of some reference sites using biological data (i.e., unexpectedly low scores at some reference sites) suggested that the original selection criteria be reassessed (i.e., made more strict). The effect of adjusting several of the reference site criteria on the number and regional distribution of reference sites was evaluated.

Section 2.2 details the process used to select 145 reference sites from the original data set. Note that a site had to meet all of the 12 reference site criteria to be identified as a reference site. The original breakdown of reference sites by strata and calibration/validation set (after excluding sites less than the minimum watershed size) is given in Table 2-3. The effect of changing the following criteria were examined both individually and in combination, with these results:

- If DO were changed to $\geq$ 5 ppm instead of $\geq$ 4 ppm, 3 of the original145 sites would be dropped, leaving a total of 142 reference sites.

- If optimal remoteness score were required instead of suboptimal-to-optimal, 59 sites would be dropped, leaving 86 sites.

- If optimal aesthetic rating were required instead of suboptimal-to-optimal, 31 sites would be dropped, leaving 114 sites.

- If optimal instream habitat score were required instead of suboptimal-to-optimal, 73 sites would be dropped, leaving 72 sites.

- If all of the above changes were instituted, 112 sites would be dropped, leaving 33 sites.

The regional breakdown that would result from this combination of adjustments is shown in Table 4-4. The number of resultant sites would clearly be too few to support IBI development for each region. Therefore, no changes in reference site criteria are recommended at this time.

However, the issue of continuing to identify the most appropriate reference sites is one of ongoing interest to the MBSS program. The MBSS plans to periodically resample a selected group of "sentinel" reference sites to monitor whether conditions there have changed. In addition, with more data collected in future MBSS sampling, it may be beneficial to revisit the definition of reference sites, given that this larger pool of sites may support the use of more stringent reference site criteria. In general, future sampling should provide more data to better characterize reference conditions and their natural biological variability.

Table 4-4. Number of reference sites that would be identified in fish IBI calibration and validation data sets under full suite of adjusted reference site criteria (sites with watersheds $\geq$ 300 acres)

	Coastal Plain	Eastern Piedmont	Highlands	Total
Calibration				
Reference	5	5	13	23
Validation				
Reference	4	2	4	10
TOTAL				33

4.5 DIRECTIONS FOR FUTURE RESEARCH

Additional information on the analysis and interpretation of fish IBI scores will be valuable as the State continues to develop new applications for this ecological indicator. The MBSS may want to undertake follow-up investigations with sampling specifically targeted to address critical issues such as interannual and within-season variability. Spatial variability, in particular the extent to which one or several sites can be used to characterize entire stream reaches or watersheds, remains another key issue. MBSS is currently conducting field sampling to examine the spatial extent of reference conditions upstream and downstream of existing reference sites. It is also examining how natural factors that may influence IBI scores can be incorporated in stream condition ratings(e.g., shallow streams may be underrated by the current IBI, even after correcting for watershed size). Beginning in the year 2000, MBSS sampling will focus on providing stream condition estimates for smaller watershed units; this should provide additional data with which to examine the robustness of IBI assessments.

Finally, further development of ecological indicators for waters not yet rated by the fish IBI will be explored. Targeted sampling in coldwater and blackwater streams would be an important first step. New data from continuous temperature loggers will provide better information for classifying coldwater streams and evaluating thermal impacts. In very small headwater streams not currently rated by the fish IBI, the benthic IBI and a streamside salamander indicator (currently under consideration) could provide valuable information to assess these important resources.

5. ACKNOWLEDGMENTS

The MBSS is a cooperative effort involving many individuals from the Maryland Department of Natural Resources, Versar, Inc., and the University of Maryland's Appalachian Laboratory and Wye Research and Education Center. This project was supported in part by contract no. PR-96-055-001 from the Maryland Department of Natural Resources' Power Plant Research Program.

6. REFERENCES

Angermeier, P.L. and J.R. Karr. 1986. Applying an index of biotic integrity based on stream-fish communities: Considerations in sampling and interpretation. *North American Journal of Fisheries Managment*. 6:418-429.

Baker, J.P., D.P. Bernard, S.W. Christensen, M.J. Sale, J. Freda, K.J. Heltcher, D.R. Marmorek, L. Rowe, P.F. Scanlon, G.W. Suter II, W.J. Warren-Hicks, and P.M. Welbourn. 1990. Biological effects of changes in surface water acid-base chemistry. Acidic Deposition: State of Science/Technology Report No. 13. National Acid Precipitation Assessment Program.

Balon, E.K. 1975. Reproductive guilds of fishes: A proposal and definition. *Journal of the Fisheries Research Board of Canada* 32:821-864.

Barbour, M.T., J. Gerritsen, G.E. Griffith, R. Frydenborg, E. McCarron, J.S. White, and M.L. Bastian. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society* 15:185-211.

Barbour, M.T., and J.B. Stribling. 1991. Use of habitat assessment in evaluating the biological integrity of stream communities. In Biological Criteria: Research and Regulation. U.S. Environmental Protection Agency, Washington, D.C. EPA-440/5-91-005. pp. 25-38.

Berkman H.E. and C.F. Rabeni. 1987. Effect of siltation on stream fish communities. *Environmental Biology of Fisheries* 18:285-294.

Boesch, D.F. 1977. Application of numerical classification in ecological investigations of water pollution. U.S. Environmental Protection Agency, Office of Research and Development, Corvallis, Oregon. EPA-600/3-77-033.

Burkett, M.J. and R.P. Morgan II. 1996. Metric selection for an index of biotic integrity: Developing an IBI for the Upper Monocacy River Drainage. Final Report to Tidewater Administration, Maryland Department of Natural Resources, Annapolis, MD.

EA Engineering, Science, and Technology, Inc. (EA). 1993. Development and modification of an index of biotic integrity for New Jersey lotic waters. EA project 60341.01. Hunt Valley, MD.

Fore, L.S., J.R. Karr, and L.L. Conquest. 1994. Statistical properties of an index of biological integrity used to evaluate water resources. *Canadian Journal of Fisheries and Aquatic Science* 51:1077-1087.

Fore, L.S., J.R. Karr, and R.W. Wisseman. 1996. Assessing invertebrate responses to human activities: Evaluating alternative approaches. *Journal of the North American Benthological Society* 15:212-231.

Hall, L.W., Jr., S.A. Fischer, W.K. Killen, Jr., M.C. Scott, M.C. Ziegenfuss, and R.D. Anderson. 1993. A pilot study to evaluate biological, physical, chemical, and land-use characteristics in Maryland coastal plain streams. Chesapeake Bay Research and Monitoring Division, Maryland Department of Natural Resources. CBRM-AD-94-1.

Hall, L.W., Jr., R.P. Morgan, E.S. Perry, and A. Waltz. 1999. Development of a physical habitat index for Maryland freshwater streams. Draft Report to Maryland Department of Natural Resources, Monitoring and Non-tidal Assessment Division, Annapolis, MD.

Hall, L.W., Jr., M.C. Scott, and W.D. Killen Jr. 1996a. Development of biological indicators based on fish assemblages in Maryland coastal plain streams. Prepared by the University of Maryland, Wye Research and Education Center, Queenstown, MD, for Maryland Department of Natural Resources, Monitoring and Non-Tidal Assessment Program, CBWP-MANTA-EA-96-1.

Hall, L.W., Jr., M.C. Scott, W.D. Killen Jr., and R.D. Anderson. 1996b. The effects of land-use characteristics and acid sensitivity on the ecological status of Maryland coastal plain streams. *Environmental Toxicology and Chemistry* 15:384-394.

Jacobson, P., P. Kazyak, A. Janicki, D. Wade, H. Wilson, and R.P. Morgan, II. 1992. Feasibility of using an Index of Biotic Integrity (IBI) approach for synthesizing data from a Maryland Biological Stream Survey. Prepared by Versar, Inc., Columbia, MD, for Maryland Department of Natural Resources, Chesapeake Bay Research and Monitoring Division. CBRM-AD-93-11.

Jenkins, R.E. and N.M. Burkhead. 1993. *Freshwater Fishes of Virginia*. American Fisheries Society, Bethesda, MD.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. *Fisheries* 6(6):21-27.

Karr. J.R. 1991. Biological integrity: A long-neglected aspect of water resource management. *Ecological Applications* 1:66-84.

Karr, J.R. and D.R. Dudley. 1981. Ecological perspectives on water quality goals. *Environmental Management* 5:55-68.

Karr, J.R., K.D. Fausch, P.L. Angermeier, P.R. Yant, and I.J. Schlosser. 1986. Assessing biological integrity in running waters: A method and its rationale. Illinois Natural History Survey Special Publication 5. 28 pp.

Kazyak, P.F. 1996. Maryland Biological Stream Survey sampling manual. Maryland Department of Natural Resources, Monitoring and Non-tidal Assessment Division.

Kazyak, P.F., R.L. Dwyer, and S.B. Weisberg. 1988. Baseline characterization of the Gwynns Falls watershed prior to nonpoint source remediation. Prepared by Versar, Inc., Columbia, MD for Maryland Department of Natural Resources, Tidewater Administration.

Knapp, C.M., W.P. Saunders, D.G. Heimbuch, H.S. Greening, and G.J. Filbin. 1988. Maryland Synoptic Stream Chemistry Survey: Estimating the number and distribution of streams affected by or at risk from acidification. Prepared by International Science and Technology, Inc., Reston, VA, for the Maryland Department of Natural Resources, Power Plant Research Program, Annapolis, MD. AD-88-2. NTIS No. PB88-213996/AS.

Lee, D.S., S. Platania, C.R. Gilbert, R. Franz, and A. Norden. 1981. A revised list of the freshwater fishes of Maryland and Delaware. *Proceedings of the Southeastern Fishes Council* 3(3):1-10.

Leidy, R.A., and P.L. Fielder. 1985. Human disturbance and patterns of fish species diversity in the San Francisco Bay drainage, California. *Biological Conservation* 33:247-267.

Leonard, P.M. and D.J. Orth. 1986. Application and testing of an index of biotic integrity in small, coolwater streams. *Transactions of the American Fisheries Society* 115:401-414.

Lyons, J., L. Wang, and T.D. Simonson. 1996. Development and validation of an index of biotic integrity for coldwater streams in Wisconsin. *North American Journal of Fisheries Management* 16:241-256.

Miller, D.L., P.M. Leonard, R.M. Hughes, J.R. Karr, P.B. Moyle, L.H. Schrader, B.A. Thompson, R.A. Daniels, K.S. Fausch, G.A. Fitzhugh, J.R. Gammon, D.B. Halliwell, P.L. Angermeier, and D.J. Orth. 1988. Regional applications of an index of biotic integrity for use in water resource management. *Fisheries* 13(5):12-20.

MRLC (Multi-Resolution Land Characteristics Consortium) 1996. U.S. Federal Region III land cover data set: Version 2.0 metadata. MRLC website, http://www.epa.gov/mrlc/R3Meta.

Mundahl, N.D. and T.P. Simon. 1999. Development and application of an Index of Biotic Integrity for coldwater streams of the Upper Midwestern United States. In T.P. Simon (ed.). *Assessing the Sustainability and Biological Integrity of Water Resources Using Fish Communities*. Boca Raton, FL: CRC Press.

Ohio EPA (Ohio Environmental Protection Agency). 1987. Biological criteria for the protection of aquatic life. Volumes I-III. Ohio EPA, Division of Water Quality Monitoring and Assessment, Surface Water Section, Columbus, Ohio.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* 77:118-125.

Osborne, L.L. and D.A. Kovacic. 1993. Riparian vegetated buffer strips in water-quality restoration and stream management. *Freshwater Biology* 29:243-258.

Page, L.M. 1983. *Handbook of Darters.* TFH Publication, Inc., Neptune City, New Jersey.

Pflieger, W.L. 1975. *The Fishes of Missouri* . Missouri Department of Conservation, Jefferson City, MO.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross and R.M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: Benthic macroinvertebrates and fish. U.S. Environmental Protection Agency, Office of Water, Washington, D.C. EPA 440-4-89-001.

Ranasinghe, J.A., L.C. Scott, and S.B. Weisberg. 1996. Chesapeake Bay Water Quality Monitoring Program: Long-term benthic monitoring and assessment component, Level I comprehensive report. Prepared by Versar, Inc., Columbia, MD, for Maryland Department of Natural Resources, Tidewater Ecosystem Assessment Division.

Rankin, E.T. 1989. The Qualitative Habitat Evaluation Index (QHEI): Rationale, methods, and application. Ohio EPA, Division of Water Quality Planning and Assessment, Ecological Analysis Section, Columbus, OH.

Rankin, E.T. 1995. Habitat indices in water resource quality assessments. In Davis, W.S. and T.P. Simon, eds. 1995. *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making.* Lewis Publishers, Boca Raton, FL.

Reger, J. 1995. Memorandum from J. Reger, Maryland Geological Survey, Environmental Geology and Mineral Resources, to J. Perdue, A. Raspberry, and E. Bradley. Digital files and metadata for draft physiographic map of Maryland. 14 July 1995.

Roth, N.E., M.T. Southerland, J.C. Chaillou, R.J. Klauda, P.F. Kazyak, S.A. Stranko, S.B. Weisberg, L.W. Hall, Jr., and R.P. Morgan II. 1998a. Maryland Biological Stream Survey: Development of a fish index of biotic integrity. *Environmental Management and Assessment* 51:89-106.

Roth, N.E., M.T. Southerland, J.C. Chaillou, H.T. Wilson, D.G. Heimbuch, and J.C. Seibel. 1998b. Maryland Biological Stream Survey: Ecological status of non-tidal streams sampled in 1996. Prepared by Versar, Inc., Columbia, MD, and Post, Buckley, Schuh, and Jernigan, Inc., Bowie, MD, for Maryland Department of Natural Resources, Monitoring and Non-Tidal Assessment Division. CBWP- MANTA-EA-98-1

Roth, N.E., M.T. Southerland, J.C. Chaillou, J.H. Volstad, S.B. Weisberg, H.T. Wilson, D.G. Heimbuch, and J.C. Seibel. 1997. Maryland Biological Stream Survey: Ecological status of non-tidal streams in six basins sampled in 1995. Prepared by Versar, Inc., Columbia, MD, and Coastal Environmental Services, Linthicum, MD for Maryland Department of Natural Resources, Monitoring and Non-Tidal Assessment Division. CBWP-MANTA-EA-97-2.

Roth, N.E., M.T. Southerland, G. Mercurio, J.C. Chaillou, D.G. Heimbuch, and J.C. Seibel. 1999. State of the Streams: 1995-1997 Maryland Biological Stream Survey results. Prepared by Versar, Inc., Columbia, MD, and Post, Buckley, Schuh, and Jernigan, Inc., Bowie, MD, for Maryland Department of Natural Resources, Monitoring and Non-Tidal Assessment Division. CBWP-MANTA-EA-99-6.

Roth, N.E., J.H. Volstad, G. Mercurio, and M.T. Southerland. 2000. Biological indicator variability and stream monitoring program integration: A Maryland case study. Prepared by Versar, Inc., Columbia, MD, for U.S. Environmental Protection Agency, Mid-Atlantic Integrated Assessment.

Simon, T.P. and J. Lyons. 1995. Application of the index of biotic integrity to evaluate water resource integrity in freshwater ecosystems. In Davis, W.S. and T.P. Simon, eds. 1995. *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making.* Lewis Publishers, Boca Raton, FL.

Smith, P.W. 1979. *The Fishes of Illinois*. University of Illinois Press, Urbana Il.

Southerland, M.T., N.E. Roth, and S.K. Patty. 2000. Freshwater biodiversity hotspots in Maryland. Proceedings of *Conservation of Biological Diversity: A Key to the Restoration o f the Chesapeake Bay Ecosystem and Beyond*. Maryland Department of Natural Resources, Annapolis. May 10-13, 1999.

Strahler, A.N. 1957. Quantitative analysis of watershed geomorphology. *Transactions of the American Geophysical Union* 38(6):913-920.

Stribling, J.B., B.K. Jessup, J.S. White, D. Boward, and M. Hurd. 1998. Development of a Benthic Index of Biotic Integrity for Maryland streams. Prepared by Tetra Tech, Inc., Owings Mills, MD and Maryland Department of Natural Resources, Monitoring and Non-Tidal Assessment Program. CBWP-MANTA-EA-98-3.

Trautman, M.B. 1981. *The Fishes of Ohio.* Ohio State University Press, Columbus.

USEPA (U.S. Environmental Protection Agency). 1987. Handbook of methods for acid deposition studies: Laboratory analysis for surface water chemistry. EPA-600/4-87/026.

USEPA. 1990. Biological criteria: National program guidance for surface waters. U.S. EPA, Office of Water Regulations and Standards. EPA-440-5-90-004.

USEPA. 1996a. Summary of state biological assessment programs for streams and wadeable rivers. U.S. EPA, Office of Policy, Planning, and Evaluation. Washington, DC. EPA/230-R-96-007.

USEPA. 1996b. Biological criteria: Technical guidance for streams and rivers. U.S. EPA, Office of Science and Technology. Washington, DC. EPA/822-B-94-001.

Vølstad, J.H., M.T. Southerland, S.B. Weisberg, H.T. Wilson, D.G. Heimbuch, and J.C. Seibel. 1996. Maryland Biological Stream Survey: Tthe 1994 demonstration project. Prepared by Versar, Inc., Columbia, MD, for Maryland Department of Natural Resources, Monitoring and Non-Tidal Assessment Division, Annapolis, MD. CBWP-MANTA-EA-95-9.

Weisberg, S.B., J.A. Ranasinghe, D.M. Dauer, L.C. Schaffner, R.J. Diaz , and J.B. Frithsen. 1997. An estuarine benthic index of biotic integrity (B-IBI) for Chesapeake Bay. *Estuaries* 20:149-158.

White, J.S. 1996 (draft). Subregionalization of the Northern Piedmont and Mid-Atlantic Coastal Plain ecoregions within Maryland. University of Maryland, College Park, MD, and Tetra Tech, Inc., Owings Mills, MD.

Yoder, C.O. and E.T. Rankin. 1995. Biological response signatures and the area of degradation value: New tools for interpreting multimetric data. In: Davis, W.S. and T.P. Simon, eds. 1995. *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*. Lewis Publishers, Boca Raton, FL.